COGNITIVE PROCESSES OF SIMILARITY AND COMBINATION IN CONCEPTUAL PRODUCT DESIGN ENGINEERING

By

Christopher Patrick McTeague

A thesis submitted to the University of Strathclyde
for the degree of
Doctor of Philosophy

Department of Design, Manufacturing and Engineering Management
Faculty of Engineering
University of Strathclyde
Glasgow, UK

2022

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Christopher McTeague   Date: *2022-08-17*

ACKNOWLEDGEMENTS

The content of this thesis is a product of a tremendous network of support, guidance and encouragement.

To my supervisors, Alex, Madeleine and Paul. I value the opportunity I've had to train as a researcher at the intersection of engineering design and psychology. Alex, thank you for convincing me that I would, in fact, really enjoy doing a PhD. Despite my many complaints over the years you were certainly right on that front. Thank you also for your guidance, support, and the high standards you set for your students. Madeleine, thank you for your patience and understanding while I got to grips with the most basic concepts of psychological science. You've kept me pointed in the right direction throughout the process and have gone above and beyond over the last few years to support me through to the end. To Paul, thanks for stepping in to get me over the line and keeping things as smooth as possible for me.

To my colleagues and mentors on the Imagine Design research project. Firstly, to the one who's had to put up with my incessant questions and distractions since we both started this journey in 2015. To Laura Hay, you've made every moment of this process a complete joy. Thank you for everything you've taught me and for being the excellent role model of a scientist that you are. To Gerard, Tijana, Laura Pidgeon, Mina, Pei-ling and Leila. I was spoiled to be able to work with you all over these years; doing research on your own isn't anywhere near as fun as having colleagues and friends to be in cahoots with. To Gerard and Tijana in particular, thanks for staying the course with me, sharing your knowledge and experience and for all the help in getting these experiments completed.

To my family. Look see! I told you I'd finish it. Thank you for everything. The path that I've taken to get here has been the result of heaps of privilege, support and love. I wouldn't have made it this far without having you all behind me.

To my friends. Both those that I knew before I started the PhD and those I've been lucky enough to call friends since. Thank you all for your support, for staying in touch even when I keep patching you to write my thesis, and for making me feel like I've always got someone to turn to when I need support.

To Sarah, you've probably seen better than anyone what this process has been like for me. I think the list of things I have to thank you for is endless, but thank you especially for your relentless encouragement, your consistent support, your saint-like patience and for humouring me every time I tell you that the thesis will be done "in six weeks". Never once did I doubt that you had my back and I'll value that forever.

Finally, to everyone who volunteered their time to take part in a study, showed genuine interest in my research and offered their help, please know that every one of you were instrumental in carrying out this research. If you ever need any help in return, please don't hesitate to get in touch.

ABSTRACT

To create new products that satisfy human needs, product design engineers use their technical, manufacturing and creative knowledge to create candidate ideas for new products and develop them into final designs that can be manufactured. All new products have some basis in prior knowledge, and so design can be viewed as a process of knowledge recombination. A variety of methods and tools have been developed to help designers produce novel, useful design concepts through combinational thinking. One way to improve these design aids is by understanding the cognitive processes involved and tailoring methods and tools to foster effective cognitive processing and overcome cognitive constraints. Yet, despite the broad acknowledgement that designers *do* combine ideas to create new ones, little is known about *how* designers combine ideas to create new ones. In particular, there is no knowledge about how designers combine *design concepts*, which are candidate ideas produced earlier in the design process.

The research presented in this thesis was conducted to model the cognitive processes involved in design concept combination and design concept similarity judgements. A deductive research approach was used to propose and test two cognitive models. The Dual-Process model of linguistic conceptual combination (Wisniewski, 1997a) was used as a basis for a cognitive model of design concept combination, and the dual-process view of similarity judgements was used as the basis of a model of design concept similarity judgements. Both models involve the same dual processes of comparison and scenario creation, and both models propose that the comparison process involves a process of alignment of structured mental representations. A series of research questions and hypotheses were proposed to test the models and a quasi-experimental research design was developed to evaluate them.

The proposed Dual-Process model of design concept similarity judgements was tested in two experiments and it was concluded that student designers make similarity judgements of pairs of early-stage, sketch-based design concepts via a single process of comparison. In the first experiment (n=11), designers were asked to rate the similarity of pairs of design concepts and provide written explanations for their numerical ratings. The responses overwhelmingly indicated that designers make similarity judgements by focusing on the common and different features of the pair, i.e., a comparison process. In a second experiment (n=35), five predictions of the Structural Alignment model of similarity judgements were tested. It was found that similarity can be predicted as a function of the common and different features of a pair of design concepts, consistent with a comparison-based model of similarity judgements. However, only four of the five predictions were supported and so the Structural Alignment model was rejected. This means that it was not possible to draw conclusions about how the comparison process occurs.

The proposed Dual-Process model of design concept combination was tested in one experiment (n=30). Student designers combined pairs of early-stage, sketch-based design concepts to create new design concepts that addressed the same brief. The proportion of combination types and their relationship with the similarity of the base concepts were measured and compared with the proposed model. Three kinds of combination were produced : (i) featural, (ii) relational and (iii) ambiguous. As the relative similarity of a pair of design concepts increases, the participants were increasingly likely to produce featural combinations and less likely to produce relational combinations. There was also evidence of a stimulus compatibility effect, a cut-off of relational combinations, and a defaulting to featural combinations. The featural and relational combinations and their relationship with similarity were consistent with the proposed model. However, the combination types were not fully accounted for. Thus, the proposed Dual-Process model does not fully capture the cognitive processes involved in design concept combination.

Overall, the initial proposal that both similarity judgments and combination of design concepts occur via the same cognitive processes was incorrect. Comparison is involved in similarity judgements and may plausibly be involved in combination, but while there is evidence of a scenario creation process in design concept combination, there is none for design concept similarity judgements. Additional hypotheses and experiments are proposed to facilitate further research into the cognitive basis of the comparison processes in both models. The research and findings were critiqued to identify the advantages, disadvantages and opportunities and recommendations for future research.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Term	Meaning
Concepts	
Concept	A generic term referring to either a category concept or a design concept
Category concept	Mental representations of classes of things (Murphy, 2004, 2010). For example, the concept of a 'cat' is a mental representation of all entities which are cats.
Design concept	In the context of product design engineering, design concepts are proposals for artefacts that do not yet exist but are being developed in order to fulfil functional requirements.
Feature-based processes and outputs	
Features	A generic term for constituent elements of concepts and design concepts.
Feature-set representation	Mental representations of concepts as sets of unrelated features.
Structured representations	Mental representations of concepts as sets of features with interconnecting internal relations. Structured representations have a 'relational structure' comprising entities, attributes, relations and (mathematical) functions.
Alignability	The degree of alignment of the mental structured representation of two concepts or design concepts.
Comparison	A computational-level description of a cognitive process that operates over feature-based representations of concepts and design concepts. Various comparison-based algorithms exist.
Structural alignment	An algorithmic-level description of cognitive processes that involves the alignment of the relational structure of two mental representations. Has been implicated in comparison, similarity judgements and conceptual combination.
Feature-based similarity	An individual's perception of similarity arising from the a process that operates on the intrinsic features of a concept.
Thematic processes and outputs	
Thematic relations	Extrinsic and complementary relations between concepts or design concepts.
Thematic relatedness	The degree of relatedness, or association strength, of the thematic relation between two concepts.
Scenario creation	A computational-level description of a cognitive process that operates on concepts with thematic relations.
Slot filling	An algorithmic-level description of cognitive processing that involves the joining of two concepts by filling a 'slot' in one concept with a 'head' in another. Proposed as an explanation for scenario creation.
Thematic similarity	An individual's perception of similarity arising from a scenario creation process.

Term	Meaning
Phenomena	
Conceptual combination	The interpretation of word pairs through a cognitive process of combination, resulting in a new concept that is based on the original pair. Sometimes referred to specifically as 'linguistic conceptual combination'
Design concept combination	The combination of a pair of design concepts to create a new design concept that addresses the same requirements as the base concepts.
Similarity judgements	The appraisal of the similarity of two (or more) concepts.
Design concept similarity judgements	The appraisal of the similarity of two (or more) design concepts.
Models	
Dual-Process model of similarity judgements	A cognitive model of similarity judgements in which similarity can arise from a *comparison* or *scenario creation* process.
Dual-Process model of *design concept* similarity judgements	A cognitive model of *design concept* similarity judgements carried out by designers in which similarity can arise from a *comparison* or *scenario creation* process.
Dual-Process model of conceptual combination	A cognitive model of conceptual combination in which individuals interpret novel word pairs by combining them to create new combined concepts. Combination can occur through a *comparison and construction* or a *scenario creation* process.
Dual-Process model of *design concept* combination	A cognitive model of *design concept combination* in which *designers* interpret novel word pairs by combining them to create new combined concepts. Combination can occur through a *comparison and construction* or a *scenario creation* process.

ABBREVIATIONS

Abb.	Meaning	Abb.	Meaning
AD	Alignable Difference	**IV**	Independent variable
BVSR	Blind variation and selective retention (model)	**LSA**	Latent Semantic Analysis
CAD	Computer Aided Design	**NC**	Non combinational
CARIN	Competition Among Relations in Nominals (model)	**ND**	Nonalignable difference
CPS	Creative problem solving (model)	**NR**	No response
DESSUA	Design Support System Using Analogy	**OD**	Ontological distance
DMEM	Design, Manufacturing and Engineering Management	**PBC**	Property Based Combination
DT	Design Task	**PDE**	Product Design Engineers
DV	Dependent variable	**PFM**	Purpose Function Means
ECCo	Embodied Conceptual Combination (model)	**RBC**	Relation Based Combination
FBS	Function Behaviour Structure	**RQ**	Research Question
HD	Hypothetico-deductive	**SA**	Structural Alignment (model)
ID	Identification	**SC**	Stimulus concepts
IPA	Interactive Property Attribution (model)	**SIAM**	Search for Ideas in Associative Memory (model)
IQR	Interquartile range	**SPSS**	Statistical Package for Social Sciences
IRR	Inter rater reliability	**VHS**	Video Home System

Note: Abb. = Abbreviation

RELEVANT PUBLICATIONS

The following list contains publications to which the researcher made a significant contribution and which are relevant to the research presented in this thesis.

I Hay, L., McTeague, C., Duffy, A.H., Pidgeon, L.M., Vuletic, T. and Grealy, M., 2017. *A systematic review of protocol studies on conceptual design cognition*. In Design Computing and Cognition'16 (pp. 135-153). Springer, Cham.

II Hay, L., Duffy, A.H., McTeague, C., Pidgeon, L.M., Vuletic, T. and Grealy, M., 2017. *A systematic review of protocol studies on conceptual design cognition: Design as search and exploration*. Design Science, 3.

III Hay, L., Duffy, A.H., McTeague, C., Pidgeon, L.M., Vuletic, T. and Grealy, M., 2017. *Towards a shared ontology: A generic classification of cognitive processes in conceptual design*. Design Science, 3.

IV McTeague, C., Duffy, A., Campbell, G., Grealy, M., Hay, L., Pidgeon, L. and Vuletic, T., 2017, August. *An exploration of design synthesis*. In ICED17, 21st International Conference on Engineering Design (pp. 279-288).

V McTeague, C.P., Duffy, A., Hay, L., Vuletic, T., Campbell, G., Choo, P.L. and Grealy, M., 2018, May. *Insights into design concept similarity judgements*. In 15th International Design Conference (pp. 2087-2098).

VI Hay, L., Duffy, A.H., Grealy, M., Tahsiri, M., McTeague, C. and Vuletic, T., 2020. *A novel systematic approach for analysing exploratory design ideation*. Journal of Engineering Design, 31(3), pp.127-149.

VII McTeague, C., Duffy, A., Hay, L., Vuletic, T., Campbell, G., and Grealy, M., To Appear. *A test of the Structural Alignment model for similarity judgements of design concepts.* In Design Computing and Cognition'22. Springer, Cham.

VIII Campbell, G., Hay, L., Duffy, A., Lyall, L., McTeague, C., Vuletic, T., and Grealy, M., To Appear. *Investigating the cognitive processes involved in design ideation using psychological tests.* In Design Computing and Cognition'22. Springer, Cham

Dissemination of the PhD research

Publications IV, V and VII disseminate research that was conducted during the PhD process.

Publication V provides a preliminary analysis of the results of Similarity Experiment 2 (Sim-P2, Section 7.1).

Publication VII provides an abridged report of the results of Similarity Experiment 3 (Sim-P3, (Section 7.2)).

Publication IV presents a model of design synthesis that evolved from discussions with Prof. Alex Duffy. The model was substantiated with a precursor to the literature review presented in Chapter 2.

Related studies

Publication VIII presents the results of a correlational study exploring the relationship between the novelty of design concepts produced through concept generation and cognitive abilities as measured by a battery of cognitive ability tests. The design concepts created by the participants in this experiment were used as the stimuli for the pairs of design concepts created in this thesis. The procedure for the concept generation study is reported in Appendix 3B.

Publication VI presents an approach for analysing exploratory design ideation. Part of the approach involves the coding of design concepts in terms of the problems and solutions explored by the designer. Concepts and materials were adopted from an early trial of this approach and were used during stimuli creation (Section 6.3) in an attempt to categorise design concepts into taxonomic hierarchies.

Publications II and III present two parts of the outputs of a systematic review of protocol studies of conceptual design. Publication I is a precursor to both publications. The review and the resulting classification of cognitive processes heavily influenced the research in this thesis. The literature sample for the systematic review was reanalysed to contribute to the review of combinational cognitive processes (Chapter 2). The cognitive process classification in publication III was adapted to define combination cognitive processes (Section 1.1.4).

1 INTRODUCTION

1.1 Background literature

1.1.1 Concept combination in conceptual Product Design Engineering

Product design engineering (PDE), or engineering design, are names for the activity concerned with the creation of physical products that satisfy human needs. The PDE process is generally described in some variation of six phases (Howard *et al.*, 2008), beginning with a phase of establishing a need or considering the market (Pugh, 1991) and ending with the implementation (manufacture and sale (Pugh, 1991)) of a product (Figure 1-1). Product design engineers (henceforth 'designers') are individuals trained in design and engineering that contribute to the PDE process predominantly in the middle four 'major design phases': analysis of task, conceptual design, embodiment design and detailed design (Howard *et al.*, 2008).

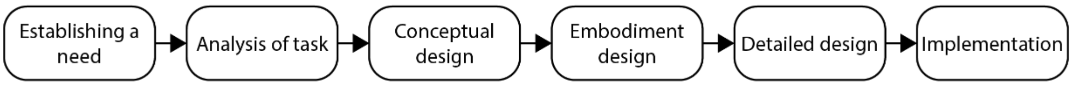


Figure 1-1 –Phases of the engineering design process (Howard *et al.*, 2008)

To contribute towards the creation of new products, designers use their technical, manufacturing and creative knowledge to translate human needs and desires into a final specification for a product that can be manufactured. To achieve this, the designers carry out a variety of design activities; rational actions taken by the designer to achieve design goals that recur across the phases of design (Sim and Duffy, 2003). Examples of these activities are the *decomposition* of problems into manageable chunks, the *specification* of requirements, the *generation* and *combination* of ideas to create candidate proposals, and the iterative development of these ideas into progressively more concrete and detailed ideas (Cash and Kreye, 2017; Sim and Duffy, 2003).

An important activity involved in design is the combination of existing knowledge to produce new ideas. This activity is important because the creation of new, novel products depends substantially on the combination or recombination of concepts and physical materials that already exist (Fleming, 2001; Nelson and Winter, 1977). Combinatorial thought also has the potential to lead to radical inventions (Schoenmakers and Duysters, 2010), those that differ from existing products or provide a basis for new directions and paradigms for technological development (Ahuja and Lampert, 2001).

Design is not the only domain in which people combine ideas to create new ones. People combine mental representations for a wide variety of purposes, including creative endeavours, meaning construction, counterfactuals, and learning and mathematics (Fauconnier and Turner, 1998). It is generally accepted that new ideas cannot be created from nothing (Finke *et al.*, 1992; Mumford *et al.*, 1997; Scott *et al.*, 2005), and so new design products must come from the designers knowledge and experiences.

The most extensive degree of combinatorial activity occurs in the conceptual design stage. In this stage, product design engineers and their colleagues iteratively create and evaluate candidate ideas (Liu *et al.*, 2003; Pugh, 1991) called design concepts. Design concepts are tentative representations of artefacts that can address functional requirements. The design concepts co-evolve alongside the formulation of a problem (Dorst and Cross, 2001; Maher, 2000; Maher *et al.*, 1996) and iterative bouts of evaluative processing (Liu *et al.*, 2003) and are developed into more mature ideas. These ideas and the decisions made during conceptual design are some of the most important in the design process (French, 1998). 75-90% of the cost of a product over its lifecycle is determined by the end of the conceptual design phase (Ullman, 2010; Wood and Agogino, 1996) and the concepts created during conceptual design set the direction and scope of all subsequent phases of the design process. Given the importance of the conceptual design stage, any improvements that can be made to the designer's creativity or efficiency can have knock-on effects on the quality and value of the resulting product.

1.1.2 Methods and tools for the support of combinational processing

To improve design practice, researchers can develop tools to help designers design more efficiently and effectively. Computer aided engineering design tools are widely used throughout the engineering and manufacture process in industry (Vuletic *et al.*, 2018). Design methods provide best-practice guidelines to help designers improve their creativity (Biskjaer *et al.*, 2017; Boeijen *et al.*, 2010) or identify strong candidate concepts during conceptual design (Frey *et al.*, 2009; Pugh, 1991).

A variety of tools and methods have been developed to help aid design combination. Manual heuristic methods (Table 1-1) prompt the designer to engage a variety of cognitive processes to help them explore variations of their design concepts. For example, the combining and building up of ideas into new ones is prescribed in Osborn's checklists (Osborn, 1957) and SCAMPER (Eberle, 1996), and morphological matrices can be used to systematically explore combinatorial variations of product configurations (Zwicky, 1967). Computational tools have been developed to present designers with stimuli that aid in analogy (Han *et al.*, 2018a; Luo *et al.*, 2021) and combination (Han *et al.*, 2018b; Luo *et al.*, 2021). For example, the Combinator provides the designer with a steady stream of stimuli to provide inspiration for new ideas (Han *et al.*, 2018b). The expert system developed by Luo *et al.* (2021) can retrieve inspirational stimuli from patent databases and manipulate the distance of the stimuli (how dissimilar or far removed they are in terms of their source domain).

Table 1-1 - Design methods that involve combination

Design method	**Method description**	**Role of combination**
Brainstorming	A method with rules and procedures for generating a large number of ideas.	Combinations and improvements of ideas are encouraged. Participants are instructed to build upon the ideas of others .
Brain writing/ Brain drawing/ 6-3-5	Various methods wherein concepts are written or drawn by an individual and passed on to another for them to build upon.	Participants are encouraged to combine generated ideas together to create new ones .
Morphological matrices (Zwicky, 1967)	An analytical and systematic approach to generating design concepts.	Sub-function concepts are combined in order to create overall solution concepts. 1 concept per sub-function is combined per principle solution .
(After) Pugh's controlled convergence matrix (Pugh, 1991)	Concepts are rated against a datum to identify stronger and weaker concepts. Pugh suggests various techniques to improve weaker concepts.	Combination is presented as a means to improve weak concepts by "bringing together parts or the whole of the existing solution" (Pugh, 1991, p. 92).
Scamper (Eberle, 1996)	A method which helps with idea creation through the use of 7 heuristics. The designer will confront a product concept with a series of questions based on these heuristics.	One of the 7 heuristics is 'combine'. Example questions: - What can be combined to improve the product? - What if the purposes of objectives of the concept(s) were combined?

Methods described in van Boeijen *et al.* (2014)

Since design takes place in the mind (Dinar et al., 2015), one way in which methods and tools can improve designer performance is by improving designer cognitive processing. Cognitive processes are the mental components involved in design activities. They can be generally defined as mental operations that act to transform mental representations (Poldrack *et al.*, 2011). Methods and tools can work to enhancing the designer's cognitive capacities or overcome limits on cognitive processing. For example, Fleming (2001) notes that "the most fundamental influence [on the process of invention] is a limitation on the number of potential components and combinations that an inventor can simultaneously consider" (p.119). This implies that combination is limited by the designers limited conceptual knowledge and the limited capacity of working memory (Baddeley, 2012). The expert system developed by Luo *et al.*, (2021) helps to overcome this limitation by increasing the designers exposure to technical concepts and allowing them to record task-relevant information. Likewise, the steady stream of stimuli provided by the combinator

(Han *et al.*, 2018b) helps to overcome limitations on the designer's ability to retrieve diverse stimuli, enabling them to work more quickly and produce more creative outcomes.

One salient characteristic of combination cognitive processing is that the similarity of the stimuli being combined influences the outcomes of the combination process. As reviewed in (Chapter 2), the similarity of pairs of concepts is related to the novelty and usefulness of the combined ideas (see Section 2.1.1.4). These findings have prompted researchers to embed means of manipulating similarity into stimuli retrieval tools (Luo *et al.*, 2021).

Despite the importance of knowledge about cognitive processing in the development of methods and tools, relatively little is known about how designers combine ideas to create new ones. Recent research has identified and classified the cognitive processes found in protocol studies of conceptual design (Hay *et al.*, 2017a, 2017b). This demonstrates that designers do combine (Daly *et al.*, 2012; Jin and Chusilp, 2006; Kruger and Cross, 2006), but provides little insight into *how* this happens. Likewise, despite the role of similarity in influencing combinational creativity, little is known about how designers make similarity judgements. To facilitate the next generation of design methods and tools, it would be advantageous to understand how designers combine knowledge and ideas, and how those processes can be augmented to improve design practice.

1.1.3 Cognitive processes, models and theories

To facilitate an investigation of designer cognitive processes, it is beneficial to provide a clear definition of what a cognitive process is and how knowledge about cognitive processes can be represented. Some foundational concepts were adopted from contemporary literature on theories and models in psychology to provide a more explicit framework for representing cognitive processes.

Theories are "bodies of knowledge that are broad in scope and aim to explain robust phenomena" (Fried, 2021, p.336). *Models* are abstractions of reality (Goel and Helms, 2014) that are narrower in scope than theories and are often more concrete and commonly represent a specific aspect of a theory (Fried, 2021). A model is defined as any "graphical, mathematical, computer-programmed, or verbal stylized representation of part of the real world, which concerns cognitive systems in interaction with their external and internal environments" (Jarecki *et al.*, 2020, p.1220). As Haslbeck *et al.* (2019) note, the relationship between theories and models is often not clear (p.3). For simplicity, the terms 'theory' and 'model' can be used interchangeably when referring to representations of human cognition in this thesis.

Theories and models represent psychological *constructs* and explain *phenomena* (Haslbeck *et al.*, 2019). Constructs are the target systems that themselves consist of interrelated components (Fried, 2021; Haslbeck *et al.*, 2019). For example, the designer's capacity to combine concepts to create new ones is a construct. Constructs (or target systems) give rise to *phenomena*, which are robust and recurring features in the world. For example, it will be seen in Chapter 2 that combining more diverse stimuli leads to more creative outcomes. These phenomena are usually evidenced by *effects* which are statistical relationships in data. For example, a statistically significant correlation between the similarity of a pair of stimuli and the novelty of the combined concepts. Effects are not necessary or sufficient for establishing phenomena (Fried, 2021), but they "more often than not help to discover and define them [and] many robust phenomena are based on robust effects (p.339).

The constructs (or target systems) under investigation in this thesis are cognitive processes. Cognitive processes can be represented by cognitive model, which can be understood in terms of a 'cognitive representational system' as adopted from Markman (2013), following Palmer (1978). The cognitive representational system has four components and can be described at three levels.

- **A represented world** – this is the domain being represented, e.g., objects, sentences, pictorial scenes or design concepts.
- **A representing world** – this is the domain that contains the representation itself, i.e., some abstraction of the represented world in the mind of the designer. The representing world loses information about the represented world but retains the information necessary to facilitate cognitive processing. There are various formalisms for describing the representing world. For example, concepts may be represented as points in geometric space, graph networks, feature sets, structured representations, or more complex mental models (Markman, 2013).
- **Representing rules** – the mapping between the represented world and the representing world. This relates to how one is supposed to understand the content of the representing world. For example, symbolic representations correspond to the represented world by semantic convention. The word 'wheel' corresponds to a person's concept of a wheel because they learned the representing rules that map words to concepts.

- **A process that uses the representations** – cognitive processes (such as combination or similarity judgements) that act on the content of the representing world.

In this thesis, *cognitive processes* are mental operations that act on and transform entities in the representing world. The interrelation between these four components is important because, as will be seen in Section 3.2.1, models make different predictions about the same constructs depending on what information is assumed to be being utilised by the cognitive process and how that information is represented.

The representing world can be described at different levels of representation (Marr, 1982; Marr and Poggio, 1976).

1. **Computational level -** Descriptions at the computational level deal with the goal(s) of the system, as well as its inputs, outputs and the relation between them.
2. **Algorithmic level** – describes how inputs are transformed into the output. This level deals with issues of representation (how are the inputs and outputs represented in the representing world) and how are those inputs transformed into the outputs.
3. **Implementational level -** refers to how algorithms are implemented in the brain.

The first two of these three levels are used throughout this thesis, but the third level is beyond the scope of this research.

Figure 1-2 shows an illustration of a framework based on the above cognitive representational system. The framework includes the represented and representing worlds. **The represented world** contains external design information such as inspirational stimuli, functional requirements, constraints and goal instructions and the outputs of whatever process is being modelled (e.g., a new design concept produced through combination). Design concepts exist in the represented world as physical or digital representations such as sketches, physical prototypes or CAD models. Designers create new design concepts through activities such as speaking, writing, sketching or modelling. **The representing world** contains the designer's mental interpretation of the represented world and their intentions for creating new content in the represented world. The designer's cognitive processes are also shown as part of the representing world. The representing world is shown at the computational and algorithmic levels. This framework is used to represent designer cognitive processes starting in Chapter 4. As Markman (2013) notes "The boundaries between the levels of description are not always sharp" (p.24), and so it is important to note that Figure 1-2 serves to provide a useful distinction between two levels but does not necessarily reflect any 'true' psychological reality.

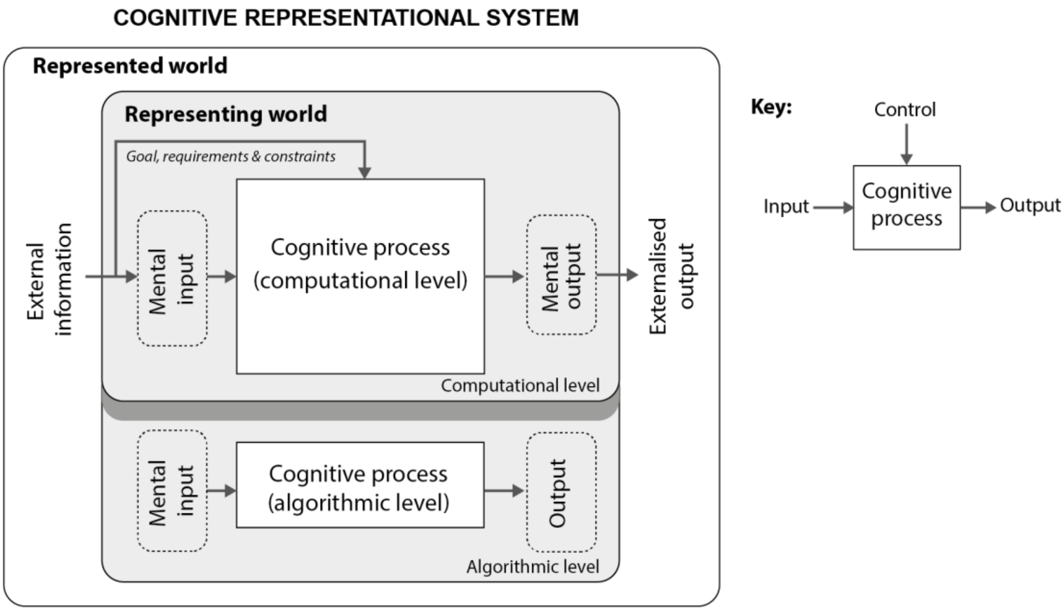


Figure 1-2 – A diagram of the cognitive representational system

1.1.4 Research focus, design concept combination

The focus of this research is on 'design concept combination', the combination of design concepts to produce new design concepts. The cognitive representational system introduced in the previous section can be used to help clarify the research focus by drawing a distinction between the process of combination and the concepts being combined.

1.1.4.1 The process of combination

The cognitive process of combination is defined as the creation of a new design entity based on two or more external and identifiable prior entities. 'Entity' is a generic term that refers to the mental representations that are processed and processed, e.g., a designers knowledge of product parts and components, sources of inspiration, previous design ideas or new design ideas. 'External and identifiable' means that the entities being combined were represented externally to the designer, such as in the form of a picture or sketch.

This definition of 'combination' is developed from an existing classification of cognitive processes (Hay *et al.*, 2017a), which can be used to help distinguish combination from other processes. This classification represents the cognitive processes involved in conceptual design, derived from a systematic review of protocol studies. One segment of this classification includes 'creative output production processes' that are involved in producing design concepts. 'Combination' (termed 'synthesis' in the original classification) may be demarcated from generating, transforming and reasoning processes (Table 1-2).

Table 1-2 – The definition of combination and three other creative output production processes

Process	**Definition**
Generating	"Producing new ideas for solutions or partial solutions to design problems"
Transforming	Using a single previously generated entity to produce a modified or entirely new entity
Combining	Using two or more external and identifiable prior entities in the creation of a new entity
Reasoning	"Thinking and drawing conclusions in accordance with some system of logic", e.g., analogical and case-based reasoning

The first distinction to be made is between combination and reasoning processes, i.e., analogical and case-based reasoning. In the context of creative output production, analogical reasoning is defined as the use of elements of solutions to problems in other domains to create new solutions to the problem that the designer is working on (Fu *et al.*, 2013). A salient characteristic of these reasoning processes is that solutions from past problem domains are mapped to the designer's current problem domain, i.e., the design problem they are currently trying to address. Combination differs from these reasoning processes in that involves the bringing together of two entities irrespective of the problems they have been used to solve in the past.

The second distinction is between 'transformation' which is defined as the creation of a new design concept based on a *single* previous idea and 'combination' which is based on *two or more* previous ideas[1]. This distinction has been made before, e.g., Daly *et al.* (2012) distinguish between 'synthesize' (merging two or more concepts) and 'elaborate' (increasing the detail of a single concept).

The third distinction is between 'generation' and 'combination'. An important distinction here is whether the entities combined are internal (i.e., retrieved knowledge) or external (i.e., perceived stimuli). Many authors define combination as the processing of previously

[1] Note that (Hay *et al.*, 2017a) do not distinguish between transforming and combining, but the distinction is useful for limiting the scope of the research in this thesis.

externalised entities. The designer has already 'generated' some new ideas through e.g., retrieval-based idea generation and then they compose (Jin and Chusilp, 2006; Sim and Duffy, 2003), assemble (Kruger, 1999; Kruger and Cross, 2006), or synthesise or merge (Daly *et al.*, 2012) two or more of them to create new ones. This form of combination is seen in outcome-based studies (Doboli *et al.*, 2014; Jang, 2014; Jang *et al.*, 2019; Nagai *et al.*, 2009), wherein participants are presented with pairs of concepts (e.g., object concepts, natural objects etc.) represented as words or pictures and are tasked with combining them to create new ones.

1.1.4.2 The concepts being combined in combination

Based on the last section, we see that combination is a generic process that acts on generic 'entities'. It is useful to make qualitative distinctions between various forms of combination.

'Design combination' refers to any kind of combinational processing that results in the production of a design concept. 'Design concepts' are the novel, tentative representations of artefacts that do not yet exist, created by designers to represent their intentions about the new product being developed. A designer could create a new design concept by combining elements of existing products, prior solutions to a given problem or inspiration drawn from nature. As long as the output is a new design concept, it may be said that design combination has occurred.

One form of design combination that has received attention in the academic literature is the combination of 'category concepts' to create new design concepts. Following Markman and Rein (2013), categories are collections of equivalent items, and 'concepts' are mental representations of those categories. This refers to peoples' semantic knowledge of the things that exist in the world. For example, there are a set of things in the world that people refer to as 'cars', and people have knowledge of the concept of a 'car' in their mind. Designers can combine category concepts to create design concepts (Section 2.1.1.2), and design support tools can present pairs of category concepts to designers to act as inspiration for combinational creativity (Han *et al.*, 2018b). As will be seen in Chapter 2, most prior research about combination in design has focused on the combination of category concepts. Importantly, however, this is not the only form of combination that occurs during conceptual design.

The focus of this thesis is 'design concept combination', i.e., the combination of design concepts to create new design concepts. This form of combination occurs when designers build upon their ideas to create new alternatives or develop them into more mature ideas. Figure 1-3 shows an example of how category concepts and design concepts can be combined in conceptual design. It is an idealised example to provide conceptual clarity that has been compiled from a student design session. The students[2] were tasked with designing a product that could carry heavy industrial equipment from an outdoor staging area to deployment site over poor terrain while reducing the risk of manual handling issues. The final product (C) is a wheeled frame with suspension and horizontal handles. The lineage of this artefact can be traced back to combinational thinking in the beginning of the conceptual design phase.

To create the product (C), ten design concepts were created by first producing ideas for design concepts that could address the brief and then *combining* those design concepts to create new ones (two of which are shown, A and B). The initial bout of combinational thought involves the combination of category concepts (semantic knowledge) to create initial design concepts. As conceptual design progressed, the students then further combined their initial design concepts to create more developed design concepts. This is the phenomena under investigation in this thesis, i.e., *design concept combination*. This example also shows how one of these design concepts (A), was highly similar to the final product, comprising the four wheeled arrangement, horizontal handles and vertical attachment members. Thus, one instance of combination led to a critical design concept that had substantial influence over the remainder of the design process.

[2] The images are extracts from a Master's level design project produced by the researcher and four fellow students in the department of Design, Manufacturing and Engineering Management at the University of Strathclyde.

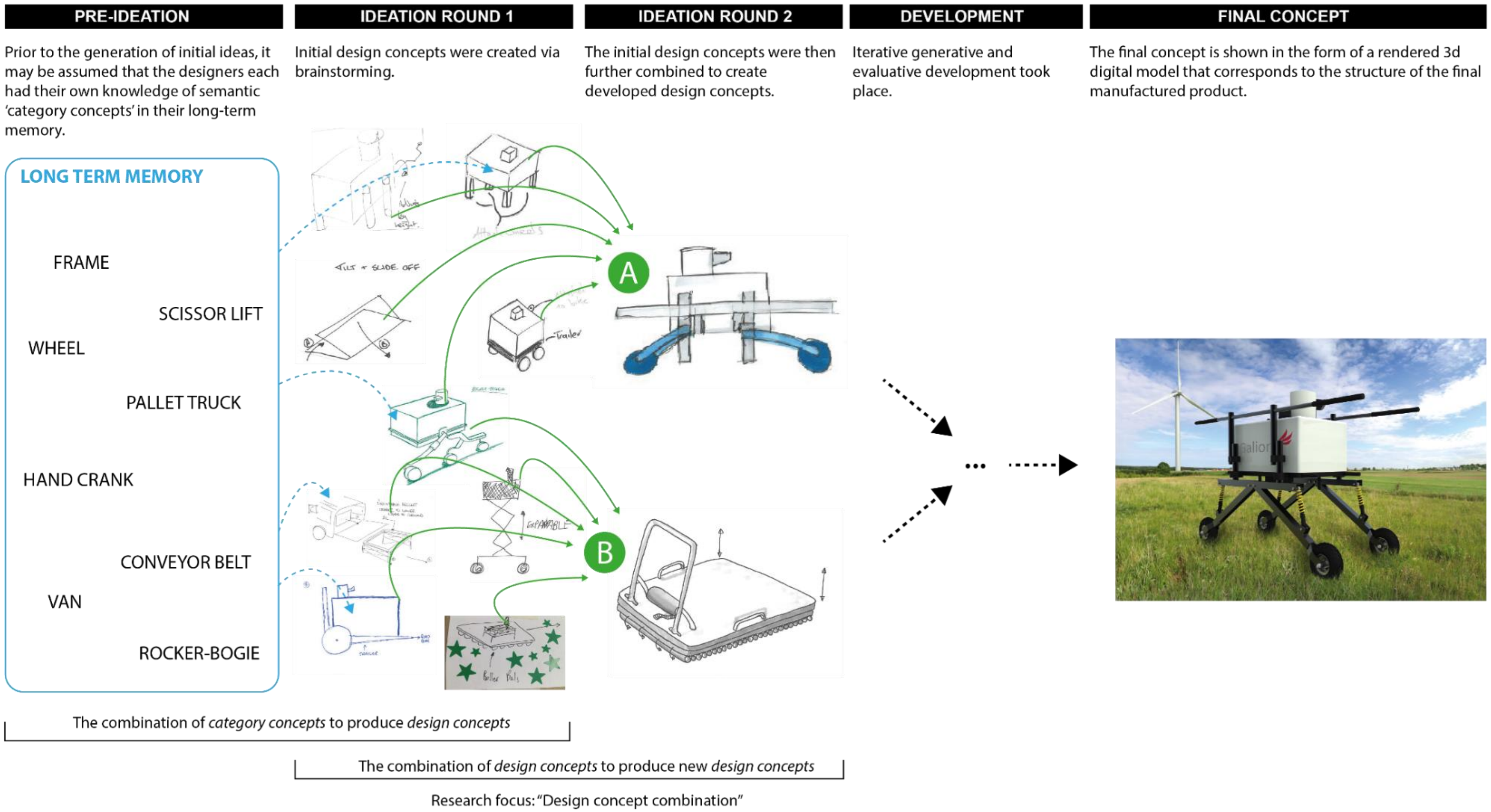


Figure 1-3 – An example of design concept combination during conceptual product design engineering

1.2 Aim and objectives

The aim of this research is to *model the cognitive process(es) of combination in conceptual product design engineering*. To achieve the research aim, the following objectives are defined as follows:

O1. Identify gaps in the current state of knowledge about combination cognitive processes and identify research methods suitable for advancing that knowledge.

- O1.1. Determine the state of knowledge about combination cognitive processes and the research methods used to study them in design.
- O1.2. Determine the state of knowledge about combination cognitive processes and the research methods used to study them in relevant non-design domains.
- O1.3. Compare the findings from design and non-design domains to identify gaps in knowledge and identify research methods suitable for advancing the current state of knowledge.

O2. Propose and test a cognitive model of design concept similarity judgements.

O3. Propose and test a cognitive model of design concept combination.

O4. Critique the work to identify strengths, weaknesses, and areas for future work.

Although the research focuses on combination cognitive processing, the investigation into combination also draws heavily from research on human similarity judgements (see Chapter 3), hence the inclusion of Objective 2.

1.3 Research approach

A research approach was developed to address the aim and objectives. The contents of the approach resemble that of the research onion (Saunders *et al.*, 2019), framed in terms of two nested components. The first component is the research philosophy, including assumptions about ontology, epistemology and axiology. The second component, which is influenced by the philosophical assumptions, is the research methodology. This includes the mode of reasoning used in knowledge creation, the steps taken to create new knowledge, and associated choices such as the intended application of the research, goodness criteria and the use of methods, techniques and data.

1.3.1 Philosophy / worldview

The term 'philosophy' is used to refer to "a system of beliefs and assumptions about the development of knowledge" (Saunders *et al.*, 2019, p.130). This generally refers to what other authors term 'paradigms' (Blessing and Chakrabarti, 2009; Guba and Lincoln, 2014) or 'worldviews' (Reich, 1994). The philosophical assumptions adopted for this thesis are explicated terms of four aspects.

- Ontology is "…concerned with the nature and relations of being (Merriam-Webster, n.d.). It deals with the nature of the world, the objects under investigation and what is meant by reality.
- Epistemology is "the study… of the nature and grounds of knowledge especially with reference to its limits and validity" (Merriam-Webster, n.d.). It deals with the development of knowledge (Wahyuni, 2012), and the question of "what can be known?" (Crossan, 2015, p.47).
- Axiology is "the study of the nature, types, and criteria of values and of value judgments especially in ethics" (Merriam-Webster, n.d.) and concerns the role of the researchers' values in the research.
- Methodology deals with the methods of knowledge creation (Reich, 1994), the modes of reasoning used to create knowledge and how it is interpreted.

There have been several attempts at mapping the landscape of the philosophical assumptions (Fleetwood, 2014; Guba and Lincoln, 2014; Ryan, 2019; Saunders *et al.*, 2019). For example, Guba and Lincoln (2014) compare four 'paradigm positions' of positivism, post-positivism, critical theory and constructivism and list their associated 'beliefs' and 'positions'. The philosophical assumptions in this thesis align with the post-positivist paradigm and there are many commonalities with transcendental/critical realism (Bhaskar, 1975). However, the use of the hypothetico-deductive method and falsificationist reasoning (Section 1.3.2.2) may make the research appear to align more closely with critical rationalism (Popper, 1959, 1963). Rather than attempting to fit the research in this thesis to any pre-existing paradigm, the philosophical assumptions underlying the thesis (Table 1-3) will be discussed in turn, stating what the assumptions are, highlighting alternatives and noting the rationale and implications for design cognition research.

Table 1-3 – Philosophical assumptions

Issue	**Implications**
Ontology	Realism.
Epistemology	Epistemological relativism – there can be multiple interpretations of reality
	Judgemental rationalism – given the existence of a single reality and the possibility for multiple interpretations of reality, it is possible to decide between interpretations
	Truth is fallible
	Modified dualist/objectivist (Guba)
Axiology	Values influence the research approach through choice of topic and desires for the creation of methods and tools in the future. Values are minimised as much as possible in the pursuit of descriptive and explanatory knowledge of designer cognitive processing
Methodology	Methodological pluralism

1.3.1.1 Ontology

Two ontological assumptions were adopted, adopted from Bhaskar's transcendental or critical realism. The two assumptions are (i) ontological realism, and (ii) reality is stratified.

(i) *Ontological realism* is the premise that there is a single, true reality that exists independently of the human mind (Pilgrim, 2020). This may be contrasted with scientific objectivism, which "claims that there is only one fully correct way in which reality can be divided up into objects, properties, and relations" (Lakoff, 1987, p.265). Scientific realism is the view that the world exists but that there can be more than one way of understanding it.

(ii) This 'true reality' is one part of a *stratified ontology* comprising three domains: the 'deep' (or 'real), the 'actual' and the 'empirical'. Observations occur in the empirical domain and are used to describe and explain the events and experiences in the 'actual' domain. The 'deep' domain involves the causal mechanisms that give rise to events and experiences in the 'actual' domain. In this thesis, observations are made in the empirical domain through quasi-experimental methods. The aim is to describe and explain the events that occur in the actual domain, such as what happens when a designer combines two design concepts to create a new one. The effects and events can only be explained with reference to the real level, i.e., the underlying causal mechanisms, but the 'reality' of the causal mechanisms that gave rise to them can only be inferred and are hidden from direct investigation.

1.3.1.2 Epistemology

The epistemological assumptions also align with critical realism, they are (i) epistemological relativism, (ii) the fallibility of truth, (iii) judgement rationalism, and (iv) modified objectivism.

(i) *Epistemological relativism* is the premise that humans have their own construals' of the world they live in and think about (Holtz and Odağ, 2020; Pilgrim, 2020). Epistemological relativism is not truth relativism (Pilgrim, 2020); knowledge is subject to different construals but truths about reality can exist independent of socials beliefs.

(ii) *Fallibilism* is notion that knowledge can be demonstrated to be incorrect. The inferences drawn from research can be wrong for a variety of reasons (e.g., type 1 and 2 errors). Even with considerable cumulative evidence, something previously accepted as 'fact' can be shown to be wrong.

(iii) *Judgemental rationalism* is the premise that given the existence of a 'real' domain (ontological realism) and accounting for alternative construals of reality (epistemological relativism), it is possible to evaluate the likelihood that different perspectives may be true and pick the best inference (Pilgrim, 2020).

(iv) *Modified objectivity* is the position that dualism (the separation of the investigator and the investigated) is not possible, but it is desirable to maximise objectivity as much as possible (Guba and Lincoln, 2014). Although knowledge is fallible and it is not possible to control or remove all social context or bias, a focus should be placed upon controlling as much external social influence as possible (Ryan, 2019).

1.3.1.3 Axiology

The axiological assumptions of the research follow on from the assumption of ontological realism. In discussing axiology, a distinction may be made between subjectivity and values. The former relates to the general influence that the researcher may exert on the subject of the research (designers). The latter refers specifically to the researchers hopes, desires, and ideas of what *ought* to be.

In line with post-positivism, the research was conducted in an attempt to eliminate the influence of researcher subjectivity as much as possible. Nonetheless, some aspects of the research involve qualitative analyses of data that are prone to bias and subjectivity. Efforts are made to reduce these biases by using measures of interrater reliability, but some subjectivity may remain. Likewise, attempts are made to limit the influence of researcher values. However, values are inherently reflected in the selection of a topic to study (Hill, 1984; Ponterotto, 2005). Although the aim of this research is to understand designer cognitive processes, this is carried out with a view towards creating interventions that can ‘improve’ design practice which innately contains some value judgements about what ‘good’ design looks like.

1.3.2 Methodology and methods

Methodology “refers to a model to conduct research within the context of a particular paradigm” (Wahyuni, 2012, p.72). Research methods, such as interviews, case studies or controlled experiments are specific procedures that can be used to collect and analyse data. The research in this thesis followed a version of the hypothetico-deductive method using quasi-experiments as the primary research method.

1.3.2.1 Selection of methodology

The methodology that was used followed on in part from the ontological, epistemology and axiological assumptions of the research (Ponterotto, 2005), as well as the epistemological assumption of methodological pluralism that no methodology is necessarily more appropriate than any other. It was also informed by the aim of developing algorithmic level models of designer cognition and examples of how the HD method can be used to create such models in other domains (see Section 2.4.3). To arrive at the selected methodology, four guiding questions were established. They were

1. What do we know about combination, i.e., what is the current state of knowledge about combination cognitive processes in design?
2. What *can* be known about combination, i.e., what kinds of knowledge can and does exist about cognitive processes more generally?
3. How can knowledge be gained about combination, i.e., what methodologies and methods can be used to create the desired knowledge in design?
4. How should that knowledge be evaluated, i.e., what constitutes ‘goodness’ or how can the worth of the work be established?

The first two questions were addressed through a literature review (Chapter 2). It was identified that there was knowledge about combination at the computational-level (Q1) but there was almost no empirically-supported knowledge about algorithmic-level processes, especially in comparison to the knowledge about combinational processes in non-design domains (Q2). In response to Q3, it was identified that new knowledge can be created by using a hypothetico-deductive methodology and adopting theories and methods from non-design domains. The answer to Q4 then followed from the choice of methodology; the research was evaluated in terms of four types of validity (see Section 4.2.5).

1.3.2.2 Hypothetico-deductive methodology

Generally, the HD method (Figure 1-4) involves proposing a theory (or model), generating hypotheses and testing them. The theory is corroborated if it is not falsified by repeated testing (Popper, 1959) and it is falsified if a reproducible effect can be discovered that refutes it (see also Fidler *et al.*, 2018). The specific implementation of the HD methodology is discussed subsequently and the methodological positions on a range of practical issues are listed in Table 1-4.

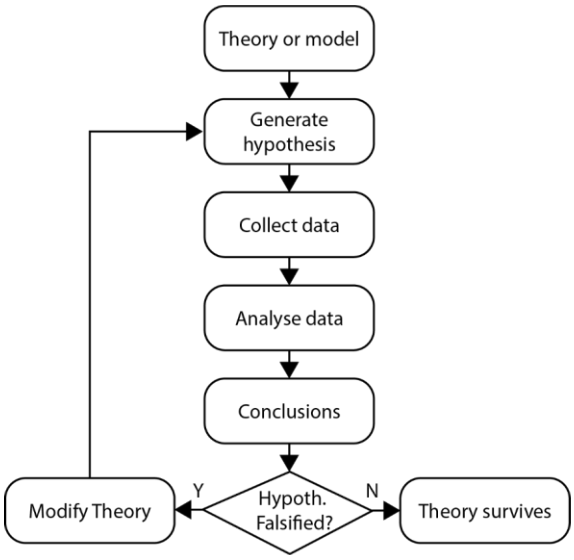


Figure 1-4 – The Hypothetico-Deductive research methodology, adapted from Spielman *et al.* (2020)

Table 1-4 – Methodological positions on practical issues

Issue	Methodological position
Inquiry aim	Description and explanation
Application potential	Basic science (as opposed to e.g., translational or applied).
Nature of knowledge	A mixture of confirmationism and falsificationism
Goodness or quality criteria	Reliability and validity (statistical conclusion, construct, internal, external)
Modes of reasoning	Abduction and deduction
Methods and data	Mixed/multi methods, predominantly quasi-experimental and quantitative data. Supplemented with interviews and qualitative data. The majority of quantitative data is derived from interpretation of qualitative data, but some data are direct numerical measures.

The aims of the research were descriptive (to describe the computational-level of design concept similarity judgements and combination) and explanatory (to explain how, for each model, the inputs were transformed into outputs). The research may be classed as basic science in that it is concerned with understanding the *what* and the *how* of designer cognition, rather than e.g., translational or applied science that would involve the testing of interventions in the lab or in practice.

A modified version of the HD method was used in this thesis (for the full research design, see Section 4.2). Cognitive models were proposed and were tested by answering research questions and testing hypotheses. The answers to research questions were assessed to determine the extent to which they provided preliminary support for the proposed models, i.e., a confirmationist approach to the HD method. Hypotheses were tested using a falsificationist approach. The goodness criteria for knowledge claims were reliability for coding qualitative data and validity for making inferences from quantitative data.

Table 1-5 – Modes of reasoning (Borsboom *et al.*, 2021)

Issue	Implications
Induction	Making generalised inferences from particular cases
Deduction	Deriving implications from general laws
Abduction	Inference to the best explanation

There are different modes of reasoning (alternatively termed approaches to theory development (Saunders *et al.*, 2019, p.152)). These include induction, deduction and abduction (Table 1-5) (but see also: Bayesianism (Fidler *et al.*, 2018) and retroduction

(Pilgrim, 2020)). The research in this thesis involved a mixture of abductive and deductive and reasoning. The methodology was broadly deductive in that it began with a model and subjected it to testing. That initial model was created through a process of analogical abduction, that is, the borrowing of explanatory principles from another domain in which a similar set of phenomena are better understood (paraphrased from Borsboom *et al.* (2021)).

1.3.2.3 Methods

The research may be described as multi-method, including a mixture of quantitative and qualitative methods. Quasi-experiments were the main research method, and non-standardised interviews and observations were used as supplementary methods.

- Correlational quasi-experiments - Experiments, quasi-experiments and correlational studies are defined in Table 1-6. The studies in the proposed research design are partly quasi-experimental because they involve the manipulation of alignability or similarity (e.g., pairs of concepts that span a range of similarity ratings) but lack the multiple conditions and random assignment typically associated with experiments. They are partly correlational because many of the predictions made by each model concern the direction of association between two variables (e.g., similarity and the number of listed alignable differences).
- Semi-structured interviews – interviews with some predetermined questions but may be followed up by unplanned questions. Used both as a means of evaluating experimental procedures and to prompt participants to introspect about their cognitive processing.
- Observations – participants were observed during some experiments so that the experimenter could identify procedural issues and time the participant as they moved through pilot studies.

Table 1-6 – Definitions of three types of study (Shadish *et al.*, 2002)

Type of study	Definition
Experiment	A study in which an intervention is deliberately introduced to observe its effects
Quasi-experiment	An experiment in which units are not assigned to conditions randomly.
Correlational study	A study that simply observes the size and direction of a relationship among variables

Early in the research process, a range of methods were considered. For example it has been said that protocol analysis is the only, or most likely, method capable of elucidating designer cognition (Cross, 2001), but there are limits to the degree to which introspection can reveal designer cognitive processes such as retrieval and perception (Lloyd *et al.*, 1995). Research methods for studying design cognition (Table A1- 1) provide a host of different approaches, but yet none appeared to provide an off-the-shelf solution for creating algorithmic-level models of designer cognitive processes.

The selection of methods was ultimately driven by the process of analogical abduction mentioned previously. Cognitive models from prior psychological studies were used to propose new cognitive models of designer cognition. Just as these models had been tested using quasi-experiments, so too were quasi-experiments adopted as the main research method in this thesis.

1.3.3 Overview of the research

Figure 1-5 shows an overview of the research presented as a series of six linear stages. In reality, these stages did not have concrete boundaries, but they serve to represent six sequential foci of the research. The figure also shows the main elements of the research and how they relate to each stage.

1. **Literature review of combination cognitive processes**
   The beginning of the research program was focused on reviewing the existing knowledge about combination cognitive processes in design and non-design domains.
2. **Definition of gaps in knowledge to be filled**
   The findings of the literature review in the two domains were compared to reveal the gaps in knowledge in design and research methods that could be used to extend that knowledge. During this phase, the aim of creating a model of designer combination cognitive processes was established. The original set of objectives corresponded to Obj's 1, 3, and 4 (Section 1.1.4.2), i.e., the objectives involved the modelling of *design concept combination;* the objective of modelling designer similarity judgements (Obj.2) was established later.
3. **Model building and research design development**
   This phase began with the goal of proposing a model of design concept combination and develop a research approach that could be used to evaluate the model. Following the model-building process, the additional objective of creating a cognitive model of *design concept similarity judgements* was introduced. The Dual-

Process model of design concept similarity judgements and the Dual-Process model of design concept combination were used as the basis of two new models in design; the Dual-Process model of *design concept* similarity judgements and the dual process model of *design concept combination*.

4. **Empirical research**

   The two models of designer cognitive processes were tested via two parallel streams of empirical research, each following the same hypothetico-deductive methodology. This involved definition of research questions and hypotheses, the development of a study design that primarily involved quasi-experiments, and three phases of stimuli creation, data collection and data analysis.

5. **Model evaluation**

   The results from the empirical studies were used to evaluate the proposed models. The validity of the inferences was discussed in terms of four types of validity: statistical conclusion, construct, internal and external. Each proposed model is evaluated with consideration of the answers to the research questions and the results of the hypothesis tests. Additionally, because both similarity judgements and combination were hypothesised to involve the same underlying processes, inferences about one construct (e.g., similarity) are also used to draw inferences about the other (e.g., combination, and vice versa).

6. **Reflection and conclusions**

   The strengths and limitations of the thesis were considered and recommendations for future work were provided.

The same stages are represented in Figure 1-6, showing where each phase is represented in the structure of the thesis. Specific objectives and a research design for the empirical research is presented in Chapter 4.

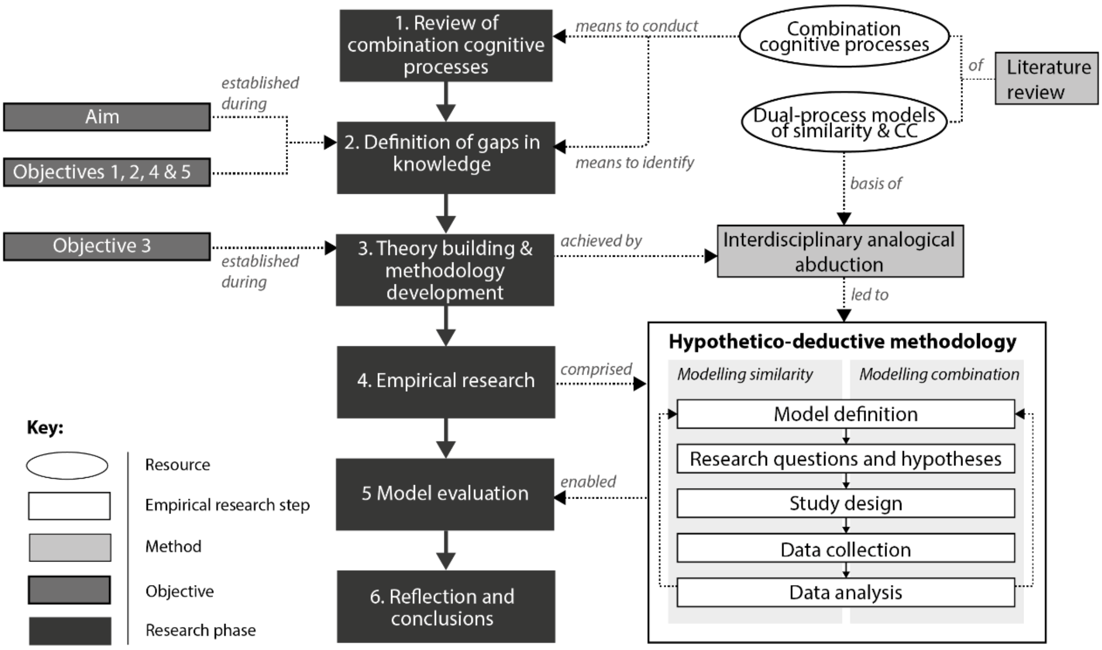


Figure 1-5 – Overview of the research. 'CC' = conceptual combination.

1.4 Structure of the thesis

Figure 1-6 provides an overview of the structure of the thesis. The 'chapters' column (grey rectangles) shows the ten thesis chapters and how they present the research that corresponds to each objective. The right side of the figure shows the content of each chapter. A distinction has been made been content associated with the design domain (blue rectangle, the majority of the outcomes in this thesis) and content in psychology and related domains (orange rectangle, representing summary outcomes from literature reviews). Dotted rectangles show the correspondence between chapters and the six phases of research shown in Figure 1-5.

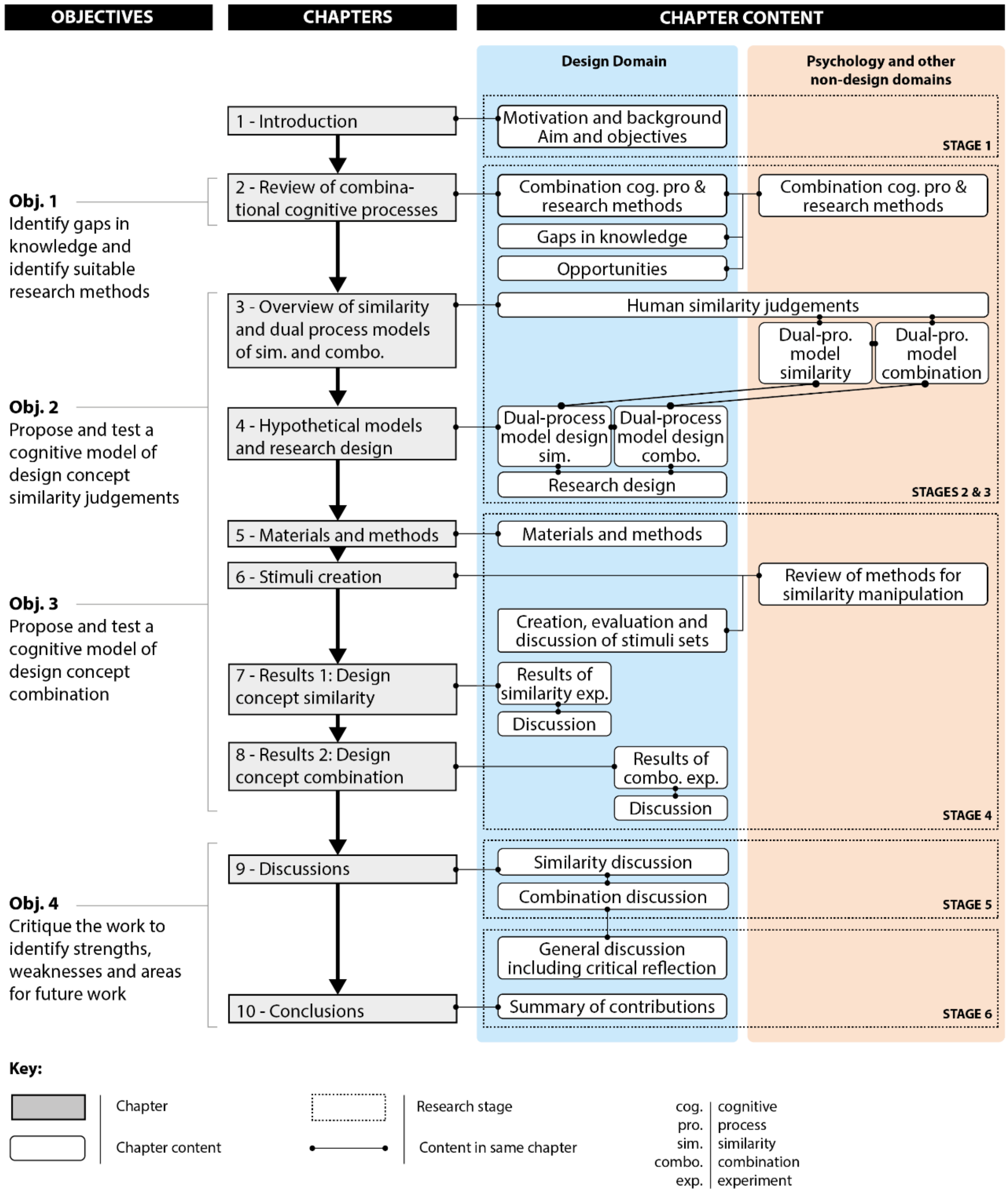


Figure 1-6 – Structure of the thesis

Literature review and gap identification (Chapter 2)

Chapter 2 presents the findings from the literature review used to address O1. Combination cognitive processes and the research methods used to study them were examined in design (O1.1) and non-design (O1.2) domains. The findings from each domain were compared to reveal the gaps in knowledge in design relative to other domains and to identify suitable research methods for advancing the current state of knowledge (O1.3).

Model building and empirical research methodology development (Chapters 3 & 4)

Chapter 3 is a second literature review chapter. It provides an overview of research on human similarity judgements generally and in design specifically. It further presents two existing cognitive models from the psychology literature, the Dual-Process model of similarity judgements and the Dual-Process model of conceptual combination.

In Chapter 4, two new models are proposed based on the existing models introduced in Chapter 3. They are the Dual-Process model of *design concept* similarity judgements the Dual-Process model of *design concept* combination. Research questions and hypotheses are defined for each model and a programme of empirical research for evaluating each model is set out. A map of the empirical research is presented, showing the chronology of the work and which chapter each element is presented in.

Empirical research (Chapters 5-8)

Chapters 5 – 8 report the empirical research that was conducted. In reality the empirical research was conducted across three sequential phases, each of which involved the creation of a set of stimuli, one or more experiments using those stimuli, and the analysis of data (Figure 4-6).

Chapter 5 presents the materials and methods and Chapter 6 presents the stimuli creation, development of coding schemes and assessment of inter-rater reliability.

Chapters 7 and 8 present the analysis, results and discussion of the experiments associated with the similarity and combination models, respectively. The discussions in Chapters 7 and 8 pertain to specific experiments, rather than the models that they were used to evaluate.

Discussions (Chapter 9)

Chapter 9 is split into three mains sections. Sections 9.1 and 9.2 evaluate the evidence for the proposed models of design concept similarity judgements and combination, respectively. The third part of the chapter (Section 9.3) is a general discussion that pertains to all of the research presented in the thesis. This includes a discussion of knowledge gained about cognitive processes in conceptual design, the research approach and the methods used. Each topic is addressed in terms of strengths, limitations and future work where appropriate.

Conclusions (Chapter 10)

Chapter 10 concludes with a summary of the knowledge contributions and the strengths and weaknesses of the research.

2 A REVIEW OF COMBINATIONAL COGNITIVE PROCESSES

'Conceptual combination' can be taken, in a general sense, to mean the mental combination of any kind of concepts such as words, visual forms, category concepts or abstract concepts such as music genres (Ward and Kolomyts, 2010). 'Design combination' is the processing of any knowledge or stimuli to produce design concepts; tentative representations for artefacts that do not yet exist. Although the extent to which human combination processes are domain-general is unknown, knowledge from non-design domains may be useful to the extent that there are measurable commonalities across various forms of combination, such as common phenomena and effects. Obj1 was thus set to assess the state of knowledge about combination cognitive processes and identify research methods from other domains that are suitable for advancing that knowledge. Three sub-objectives were set to: assess the state of knowledge about combinational processes in design (Obj. 1.1) and non-design domains (Obj. 1.2), and to compare across domains to identify gaps in knowledge in design and opportunities for future research (Obj. 1.3).

To address Obj. 1.1, a literature review was conducted focusing on combinational cognitive processes in empirical research and cognitive models of conceptual design. 'Design' was taken to mean the production of design concepts for physical artefacts or services. This includes instances of design that could be viewed as product design (Nagai *et al.*, 2009), electronic embedded systems design (Doboli *et al.*, 2014), or service design (Chan and Schunn, 2015), but excludes any research on architectural design. Based on the definition in (Section 1.1.4), design combination was defined as the creation of new design concepts based on two or more external and identifiable prior entities. Empirical research was taken from a variety of research methods, including protocol studies and experimental research (Section 2.1).

To address Obj. 2.2, a literature review was conducted on the topics of creativity and linguistic conceptual combination. Creativity refers to research on the *generic* cognitive processes that lead to the production of novel and useful outcomes, exclusive of domain-specific research on design, or other endeavours such as music (Deliège and Wiggins, 2006) or art (Chemi *et al.*, 2015). 'Linguistic conceptual combination' refers to the communicative process of interpreting novel meaning from pairs of words. They were selected based on two criteria. First, the review was conducted to identify knowledge about algorithmic-level cognitive processes that could provide some explanation for computational behaviour. This excluded highly general frameworks like the conceptual blending framework (Fauconnier and Turner, 1998) (see Section 9.3.4.1 for a discussion). The second criterion was the tasks being investigated should involve the processing of semantic or visuo-spatial knowledge to produce novel semantic concepts. These criteria led to the creativity and conceptual combination research domains, but there may be other relevant domains that were not discovered during the review process.

The literature review is presented according to three domains: design, creativity and conceptual combination. Research on conceptual combination is relatively easy to demarcate given that it is a communicative process, but as both design and creativity involve combination to produce new ideas, the boundary between the two 'domains' is less clear. Creativity is considered to be a distinct domain of research for the purposes of structuring a literature review, but as creativity is also involved in design, this boundary is partly artificial. For example, Verstijnen *et al.*, (1998a) compared design and psychology students on a task that involved combining geometric shapes to make 'objects', which were essentially design concepts. Where possible, results derived from people with design expertise were included in the 'design' section and results from those without expertise were considered in the 'creativity' section (see Table 2-7 for a summary of this chapter by domain).

To identify gaps in knowledge in design and opportunities for future research (Obj. 1.3), the findings from the design and non-design domains were compared at the computational and algorithmic levels of representation (Section 2.4.1). highlighted gaps in knowledge (Section 2.4.1.2) in the design domain and identified opportunities for new research directions in design (Section 2.4.3). Based on the results of the literature review, the Dual-Process model of conceptual combination is identified as basis for a cognitive model of design concept combination (Section 2.5).

2.1 Cognitive processes in design concept combination

Obj. 1.1 was to assess the state of knowledge about combination cognitive processes in design. Knowledge about designer cognitive processes can be gained from a variety of research methods. This makes it challenging to synthesise knowledge about designer cognitive processes. In reviewing the literature, three broad clusters of research were identified: protocol studies, outcome-based studies, and cognitive models and frameworks (listed below). To synthesise knowledge from these different studies, the cognitive representational system introduced in Section 1.1.3 was used as a common framework into which knowledge could be situated. Articles that contain knowledge about designer combination processes were identified and are listed in tables in Appendix 2A. For each representation of a combination cognitive process, the respective table lists computational-level knowledge associated with the goals, inputs and outputs of combination. The next section (2.1.1) provides a summative overview of computational level knowledge about combination cognitive processing in design.

- **Protocol studies** (Table A2-1). In protocol studies, designer cognitive processes are represented by protocol codes that are used to describe segments of designer protocol (such as think aloud verbalisations and video recordings). Individual cognitive processes are represented by qualitative labels such as 'combination' or similar terms such as 'composition' or 'assembly'. The labels may be described by written descriptions or defined in terms of abstract variables belonging to a theoretical framework. Relevant research was identified through a systematised review, in which all of the articles included in the systematic review of protocol studies of conceptual design (Hay *et al.*, 2017b, 2017a) were examined to look for combination cognitive processes.
- **Outcome-based studies** (Table A2-2). 'Outcome based' studies refer to those involving measures of independent variables (2 or more stimuli) and dependent variables (the design concepts produced through combination). In these studies, cognitive processes are the means by which the inputs (stimuli) are transformed

into outputs. Relevant research was identified through a snowball style search for empirical studies of combination processes in design.

- **Cognitive models and frameworks** (Table A2-3). The third kind of representation are cognitive models and frameworks. Although all representations of cognitive processes could be described as models (e.g., protocol coding schemes), this kind of representation are either self-titled as 'cognitive' or have some multi-component representation of verbal or graphical elements. Relevant research was identified through a snowball style search for 'cognitive models' of design which was then re-analysed to look for combinational cognitive processes.

2.1.1 Computational level knowledge: a framework of design combination variables

The computational level of combination concerns the goals of combination (what the designer is trying to achieve), the inputs to combination (the entities being combined) and the outputs (the newly created entities). Computational level knowledge is descriptive, providing an overview of the different characteristics of combination cognitive processes. Combination occurs when a designer takes two or more entities and produces a new design concept that has identifiable elements of two or more of the base concepts. Depending on what is being combined and what is being created, one may be able to identify multiple, qualitatively distinct kinds of combination. Researchers may also measure various properties of the inputs and outputs to investigate a host of factors, such as factors that contribute to creativity. Figure 2-1 shows the variables that have been used to define and measure combination in empirical studies of design cognition. Definitions of each variable are listed in tables in the relevant subsections of Section 2.1.1.

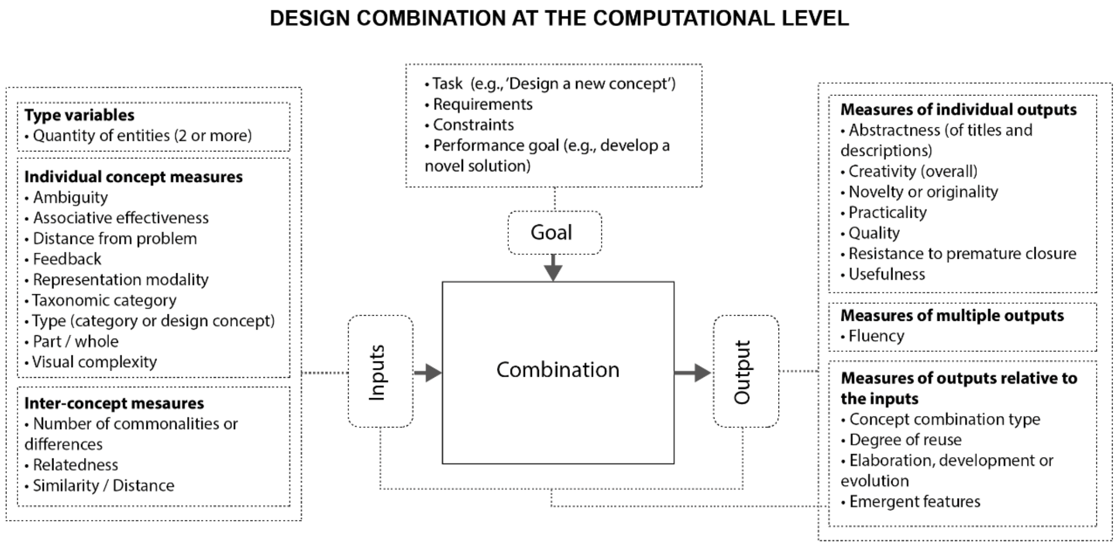


Figure 2-1 – Measures of design concept combination at the computational level

2.1.1.1 Goals

Design may be viewed as a goal-directed process (Gero, 1990; Roozenburg and Eekels, 1995), where the designer is attempting to produce a description of a design artefact to achieve requirements while adhering to constraints. The designer decomposes the overall goal into sub-goals. For example, towards the goal of (i) arriving at a single concept to take forward from the conceptual design to embodiment design phase (see: Howard *et al.* (2008)) the designer may set the goal of (ii) expanding the space of candidate solutions by (iii) combining existing concepts to create new ones.

The designer's 'top-level' goal is derived from design requirements, given to them in practice by e.g., a client or project manager. In both protocol studies and experiments designers are given design briefs to address. They may be given open-ended, problem-oriented briefs (Sosa, 2018) that define a problem to be addressed such as "How might we increase the number of registered bone marrow donors to help save more lives?" (Chan, 2014, p.19), or solution-focused briefs that specify the product to be created such as "Design an alarm clock for individual use that will not disturb others" (Glier *et al.*, 2014, p.3). Alternatively, they may be given generic instructions such as to 'design idea sketches' (Jang *et al.*, 2019) in absence of requirements. The designer's decomposed goals may differ at different phases of the design process. This does not manifest in most empirical studies of combination processes as both protocol studies and lab experiments are conducted in a single stage.

Designers may be given constraints or derive them themselves during the course of design. Daly *et al.* (2012) observed that designers choose to emphasise certain constraints and requirements, prioritising e.g., cost or portability. Alternatively, the designer may be engaging in unconstrained ideation to create a range of candidate solutions without being prematurely inhibited by thinking about constraints. None of the outcome-based combination studies (Table A2-2) gave the participants constraints over and above the requirements in the brief. Finally, in controlled experiments designers may be given performance goals such as to create novel or useful outcomes (Doboli *et al.*, 2014).

Table 2-1 - Goal variables from design combination

Variable	**Types / description**
Task	Instructions given by experimenter, e.g., 'design a new concept'
Requirements	Specification of necessary outputs. Can be to address a problem (problem-focused briefs) or to exhibit a specific function (solution-focused briefs).
Constraints	Limits on the parameters of the output
Performance goal	Instructions to maximise certain aspects of the outputs, such as "a novel electronic embedded system that is useful, use as many devices from the list as possible" (Doboli *et al.*, 2014)

2.1.1.2 The inputs of combination

The inputs of combination may be termed the 'base concepts', i.e., the entities that contain the information used in the creation of a design concept. A designer's ability to combine entities together is flexible in that they can combine a variety of different entities to create new outcomes.

Designers can combine different types of entity, including category concepts such as a guitar, ship, frog or snow (Taura, 2016, p.59), or 'abstract concepts' which are abstracted properties such as attributes or functions that can be shared by a class of entity concepts, or previously externalised design concepts. A designer may combine 'whole' concepts or elements of them. In the example in Figure 2-2, the designer combines two alternative solutions to create a new combined concept. In other cases combination may involve the composition of sub-function solutions which can be facilitated by methods such as morphological combination (Zwicky, 1967). Combination may also happen at e.g., the function, behaviour or structure levels of representation (Gero, 2000). Some processes explicitly involve the evolution, maturation or composition of elements. (Jin and Chusilp, 2006) specify that the 'compose' activity leads to more mature ideas but the nature of this maturity does not appear to be specified.

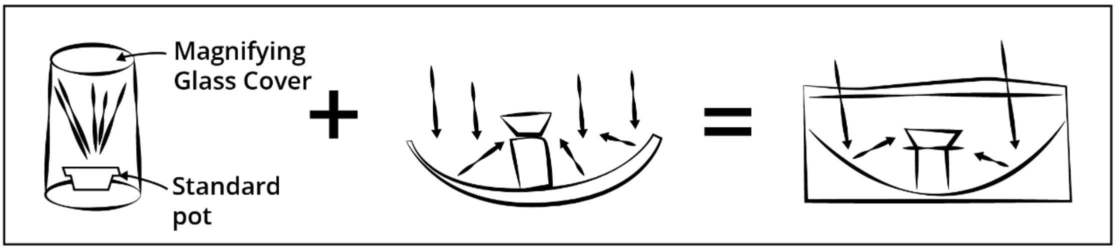


Figure 2-2 – An example of design concept combination from Daly *et al.* (2012) that appears to involve the transfer of features between two concepts.

During the course of design, one might combine a mixture of concepts (presented externally as inspirational stimuli) and design concepts (i.e., prior solutions created by the designer or a colleague). Figure 2-3 is adapted from Chan and Schunn (2015), and shows how inspirational stimuli (dotted circles) and design concepts (solid circles) are used across multiple generations of ideas towards the creation of the final concept. The number of entities may also vary. Figure 2-4 is adapted from Gonçalves and Cash (2021) and shows a sequence of 13 'nodes' representing consecutive ideas created during a design session. The arcs between ideas represent explicit links between ideas, meaning the designer provided an overt indication that the two ideas were related. For example, the links may reflect the designer being 'reminded of' an earlier idea or having gestured towards a previous idea. The thick black arcs represent a combinatorial idea, which is one that connects to many previous ideas from a session. Note, however, that the arcs do not necessarily imply that the combinatorial idea contains identifiable components or features from the prior ideas.

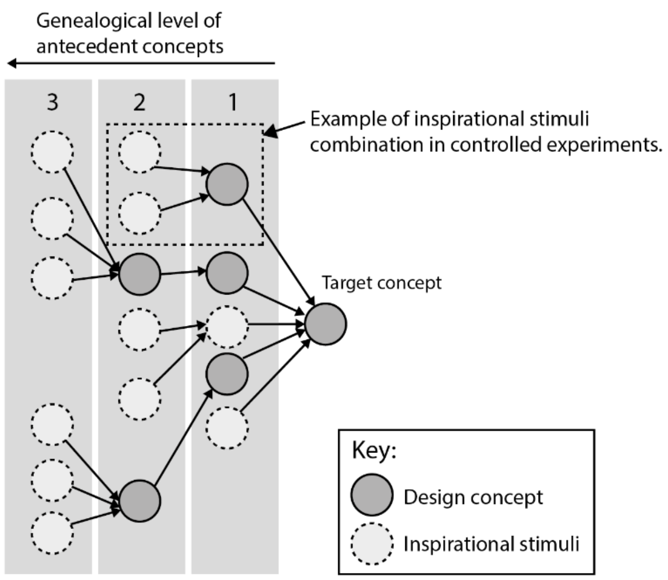


Figure 2-3 –Example of the genealogy of antecedent entities relative to a single target concept. Modified from (Chan and Schunn, 2015)

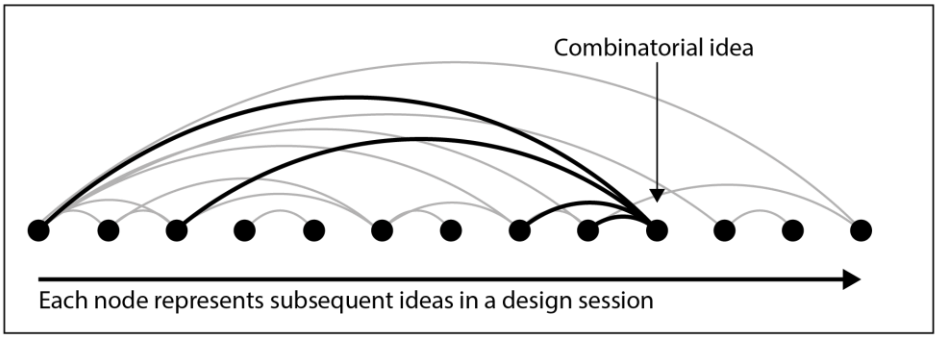


Figure 2-4 – A sequence of ideas showing instances of combination. Adapated from Gonçalves and Cash (2021)

A variety of experimental measures of individual design concepts have been taken to investigate the effects of those variables on outcome measures. These are: ambiguity of representation, associative effectiveness, distance from problem, the number of comments given as feedback, the number of commonalities and differences, representation modality, taxonomic category, type of concept (whether a design concept or category concept) and visual complexity. More individual concept properties may be investigated in the future, such as the variables studied in inspirational stimuli experiments (Vasconcelos and Crilly, 2016).

A unique aspect of combination is that properties of multiple concepts (such as pairs) can influence combination outcomes. These include measures such as distance, similarity, and relatedness.

Table 2-2 - Input variables to design combination

Variable	Types / description	Source
Number of concepts		
Quantity of entities	Quantity of distinct stimuli used as inputs to combination.	(Doboli *et al.*, 2014)
Individual concept variables		
Ambiguity	Pictorial line drawings were made more ambiguous by blending them with pictorial representations of different objects. Ambiguity operationalised as the percentage of people that could correctly identify the target image. Three levels of ambiguity: Vague (<55%), ambiguous (55-80%) or definite stimuli (>80%).	(Jang *et al.*, 2019)
Associative effectiveness	The number of associations for a word listed in the associative concept dictionary.	(Nagai *et al.*, 2009)
Distance from problem	Generally, a measure of how similar a source of inspiration is to the problem domain. Specifically, the mean of the reverse cosines between cited inspirations and the problem, using Latent Dirichlet Allocation.	(Chan and Schunn, 2015)
Feedback	Feedback about the concept in the form of comments from other people. Quantified as the number of comments received by a given concept.	(Chan and Schunn, 2015)
Representation modality	The medium of representation of an entity in the represented world. May be pictorial (Jang, 2014; Jang *et al.*, 2019), or text (Doboli *et al.*, 2014; Nagai *et al.*, 2009)	Multiple
Taxonomic category	Entities may be e.g., natural or artificial objects. They may be members of categories, e.g., tools, weapons, shelters, transportation. Multiple stimuli may be members of the same or different taxonomic categories.	(Jang, 2014)
Type (Category concept or design solution)	Refers to whether a concept represents a concept that represents a real-world entity that exists, or a solution that has been created in response to the active design task.	(Chan and Schunn, 2015)
Part / whole	Whether the entity is part of a product or a whole product. For example, designers can combine electronic modules (parts) (Doboli *et al.*, 2014) or whole concepts (Nagai *et al.*, 2009).	Multiple
Visual complexity	As determined by (Snodgrass and Vanderwart, 1980) normalised pictures	(Jang *et al.*, 2019)
Inter concept variables		
Number of commonalities and differences	The number of commonalities or differences that participants list when asked to compare two words and "list the common (similarities) and different features (dissimilarities) between the two".	(Nagai *et al.*, 2009)
Relatedness	Frequency of use together in solutions. "For example, GPS are mainly utilized for mobile applications, while cooking stoves and hair driers are static devices" (p.86).	(Doboli *et al.*, 2014)
Similarity or distance	Operationalised as taxonomic category membership (Jang, 2014; Jang *et al.*, 2019), or measured with human ratings on a likert-type scale (Jang, 2014). Semantic distance can be measured with latent Dirichlet Allocation (Chan and Schunn, 2015).	Multiple

2.1.1.3 The outputs of combination

The design concepts created through combination may vary in a range of properties and measurement of these properties can provide insights into how combination might be improved or the underlying processes that were involved in combination. The outcomes of combination can be measured for various purposes, such as to investigate the factors that influence creativity or other performance metrics. Additionally, by comparing the outputs of combination with the inputs it is possible to make inferences about the processes that transformed the inputs into the outputs.

Measures of combination outcomes may be taken from individual concepts, across multiple concepts, or with respect to the inputs. Individual concept measures identified in the literature were abstractness (of titles and descriptions), creativity, novelty or originality, practicality, quality, resistance to premature closure and usefulness (Table 2-3). An additional performance measure, fluency, may be taken across multiple outputs.

The design concepts created through combination may be characterised by how they have changed compared to the base concepts. Nagai *et al.* (2009) classifies design concepts into three kinds, each of which is said to be indicative of a different combination process. Degree of reuse and emergent features are both similar measures of the commonalities and differences of the outputs compared with the base concepts. Degree of reuse provides a measure of the extent to which the combined concept is comprised of elements from the base concepts. Measures of emergence refer to features of the combined concept not present in either base concept. The measure of elaboration used by Jang *et al.* (2019) is adopted from the Torrance Test of Creative Thinking and represents the 'number of added ideas', specifically decoration, colour, brightness and contrast, modification and the elaboration of a title.

Table 2-3 - Output variables from design combination

Variable	Types / description	Source
Variables for individual outputs		
Abstractness (of titles and descriptions)	Appears to refer to the extent to which written text provides additional information beyond the content of the sketch.	(Jang *et al.*, 2019)
Creativity (overall)	A compound construct comprising multiple individual components.	(Jang, 2014)
Novelty / originality	Originality: "whether the idea was original and novel" (Nagai *et al.*, 2009, p.661) Novelty: "uniqueness compared to the solutions of your colleagues and designs discussed in text- books, media, web"(Doboli *et al.*, 2014, p.86)	(Doboli *et al.*, 2014; Nagai *et al.*, 2009)
Practicality	Practicality: "whether the idea seemed achievable and feasible" (Nagai *et al.*, 2009, p.611)	(Nagai *et al.*, 2009)
Quality	A construct comprising five factors: size, cost, processing precision, power consumption, and easiness of interfacing.	(Doboli *et al.*, 2014)
Resistance to premature closure	"The degree of psychological openness" (Jang *et al.*, 2019, p.74)	(Jang *et al.*, 2019)
Usefulness	The extent to which a design concept solves a problem, comprising five constructs: "(1) Does the design satisfy a need? (2) Does it follow the design constraint? (3) Is there something similar available? (4) Is the pro- posed design better than the similar options? (5) Can it be used to build other things?" (Doboli *et al.*, 2014, p.93)	(Doboli *et al.*, 2014)
Variables for multiple outputs		
Fluency	Number of relevant idea sketches	(Jang *et al.*, 2019)
Variables of outputs relative to inputs		
Concept combination type	A classification of the outcome with respect to the inputs.	(Doboli *et al.*, 2014; Nagai *et al.*, 2009)
Elaboration (development or evolution).	Whether a concept has become more mature or evolved. Optionally, the degree of this maturation or evolution.	(Jang *et al.*, 2019)
Emergent features	Features present in the output concept that are not present in the base concepts.	(Nagai *et al.*, 2009)

2.1.1.4 The phenomena and effects of combination

Effects in combination research are statistical relationships between independent variables (the base concepts) and dependent variables (the combined concepts). Effects are related to the algorithmic level, in that any algorithm that explains how inputs are transformed into outputs should also provide a mechanistic explanation for the observed effects. The most relevant effects for combination studies are those specific to the presence of two or more base concepts, such as the effects of similarity or distance on creativity. Some authors measure the effects of individual-concept properties on combination, e.g., Jang (2014) investigates the influence of representation modality and representational ambiguity on creative outcome measures. However, these variables have been shown to influence ideation with the presentation of external stimuli more generally (Vasconcelos and Crilly, 2016) and are not considered in detail in this thesis.

Similarity (or distance) and creativity – as summarised by Chan and Schunn (2015) there has been mixed evidence for the effect of combination distance on creativity. Theoretically, more distant or dissimilar combinations are said to improve novelty because they increase the likelihood that emergent properties will be introduced. These emergent properties may arise through analogical mechanisms or from thinking about abstract, metaphorical aspects of the concepts being combined (Chan and Schunn, 2015). Across multiple domains of research, 'far' combinations (referring generally to concepts that are semantically distant, dissimilar or unrelated) have a consistent positive effect on novelty, but inconsistent effects on quality and overall creativity. This is generally reflected in the design studies.

- In electronic systems design less related exemplars lead to more novel concepts, but quality and usefulness were not seen to vary with relatedness (Doboli *et al.*, 2014).
- When designers produce blending combinations, there is a positive association between originality and the number of nonalignable differences (features present in one concept but not in another) (Nagai *et al.*, 2009).
- Combination of concepts from different categories (dissimilar) is also seen to lead to higher scores of elaboration in comparison with concepts from the same category (similar) (Jang *et al.*, 2019).
- Contradictorily, when designers are tasked with combining pairs of objects represented pictorially, overall creativity improves when stimuli are from different taxonomic categories but have high relative similarity, but similarity was not correlated with measures of novelty, 'resolution' or 'elaboration and synthesis' (Jang, 2014).

Chan and Schunn (2015) found that less similar combinations contribute to improved creativity when there is a genealogical lag (i.e., it takes time via iteration to see the benefits of dissimilar stimuli). They interpret this to mean that the inconsistent findings regarding novelty and utility in the literature occur because while single-instance of combination may produce novel concepts their immediate utility is uncertain, and iteration is needed to develop those concepts into ones with good overall creativity.

Similarity and combination type – Some studies, influenced by research on conceptual combination (Section 2.3), classify the outputs of combination into types and measure the frequency of occurrence of each type (Doboli *et al.*, 2014; Nagai *et al.*, 2009). In electronic systems design, relational combinations, where individual components are joined via external relations are the most common and property transfer relations, where features of one component are transferred to another to create a new component, are relatively infrequent (Doboli *et al.*, 2014). Nagai et al. (2009) examine relationships between combination type and the kinds of commonalities and differences mentioned in feature listing and comparison tasks. There have been too few of these studies to establish any stable phenomena from.

Emergence - The combination process can lead to the emergence of new features, i.e. features unique to the outcome concept and not found in either of the inputs. Nagai et al. (2009) found a non-linear relationship between originality and emergence. Emergent features have been proposed as important sources of novelty (Chan and Schunn, 2015) or creativity in combination (Estes and Ward, 2002).

2.1.2 Algorithmic level knowledge

Algorithmic level knowledge should provide explanations for a designer's capacity to combine concepts and in doing so should explain how effects arise, such as the relationship between base concept similarity and outcome novelty. Algorithmic-level processing representations are presented here and critiqued in terms of their capacity to explain effects.

Some authors have proposed algorithmic-level representations of ideation that have no empirical support. In the 'assembly' process (Kruger, 1999; Kruger and Cross, 2006), each decomposed process is represented by an 'inference model' which describes the fine-grain process steps according to an expert system model (Kruger, 1999). These steps are consistency check, combination and repair. Although the assembly process was identified in protocol, the more granular elements of the inference model were not used in the coding scheme and remain purely theoretical. In the model of mental iteration (Jin and Chusilp, 2006), the 'compose' activity involves the sub-activities: 'associate' and 'transform'. These processes are adapted from earlier work (Benami and Jin, 2002; Jin and

Benami, 2010) which in turn draws from the Geneplore model of creative cognition (Finke *et al.*, 1992). Again, there does not appear to be any empirical evidence for these sub-processes. The information processing model (Kim *et al.*, 2010) is an adaptation of two cognitive models of memory (Baddeley *et al.*, 2009; Shiffrin and Atkinson, 1971). The model is used to situate six 'cognitive operations' against a theoretical backdrop. The protocol data demonstrate the frequencies of these cognitive operations but no other components of the model.

In the systematised theory of concept generation (Taura and Nagai, 2013a, 2013b), there are three processes for concept generation, each that leads to a different type of combined concept and each described via a series of steps (Table 2-4). The three processes (property mapping, concept blending and thematic integration) are derived mainly from the Dual-Process model of conceptual combination (Wisniewski, 1997a) which also involves three combination processes. Nagai *et al.* (2009) show that design concepts can be coded according to the three output types, but there does not appear to be any empirical support for the proposition that the three outputs come from three distinct processes or for the steps involved in each process.

Table 2-4 – Three methods for concept generation (Nagai *et al.*, 2009).

Process	**Description**	**Type of output**	**Steps (in order of operation)**
Metaphor / Property mapping	New concepts are created by abstracting a property from an existing entity concept and mapping it to a category of existing concepts to create a new entity. For example, a 'flying car' could be created by abstracting the property of 'flying' from the category of 'bird' and mapping it to the category of 'car' to create a new entity in the class of cars.	The resulting design concept resembles one of the base concepts with a feature from the other base concept.	Select concepts Abstract properties Property mapping & concretisation
Concept blending	New concepts are created through the combination and concretisation of two abstracted properties.	The resulting design concept has abstract properties of both base concepts but is itself neither of the base concepts	Select concepts Abstract properties Blend & concretise
Thematic integration	This is the creation of a new scene or scenario in which two entities are related via a thematic relation and the creation of a new entity inspired by that scenario (elaboration).	The two base concepts are related by an external relation.	Select concepts Integrate them (Optional) Elaboration

In contrast with the previous models, Model-L (Liikkanen, 2010; Liikkanen and Perttula, 2010) provides empirical support for algorithmic-level processes involved in idea generation but does not explicitly acknowledge combination as being a distinct process from generation. The micro-level idea generation component of Model-L (Figure 2-5) is based on a psychological model of idea generation called Search for Ideas in Associative Memory (SIAM) (Nijstad and Sroebe, 2006). In SIAM and Model-L, idea generation is modelled in two phases of retrieval and production. In the retrieval phase, search cues (composed from the problem definition or previous ideas) facilitate the retrieval of 'memory representations' from long-term memory, where the association strength between the cue and the memory representations determines which representation is retrieved. In the idea production phase, the cue and retrieved memory representations are turned into ideas through rapid and generally effortless process where they are evaluated for newness before being externalised or abandoned to return to the memory search stage. Model-L does not include a distinct combination process, since "the complete solution is always a combination of solutions" (Liikkanen, 2010, p.67). Existing concepts or ideas can, however, be used as cues for the idea retrieval phase of idea generation, thereby providing a cognitive mechanism for design concept combination.

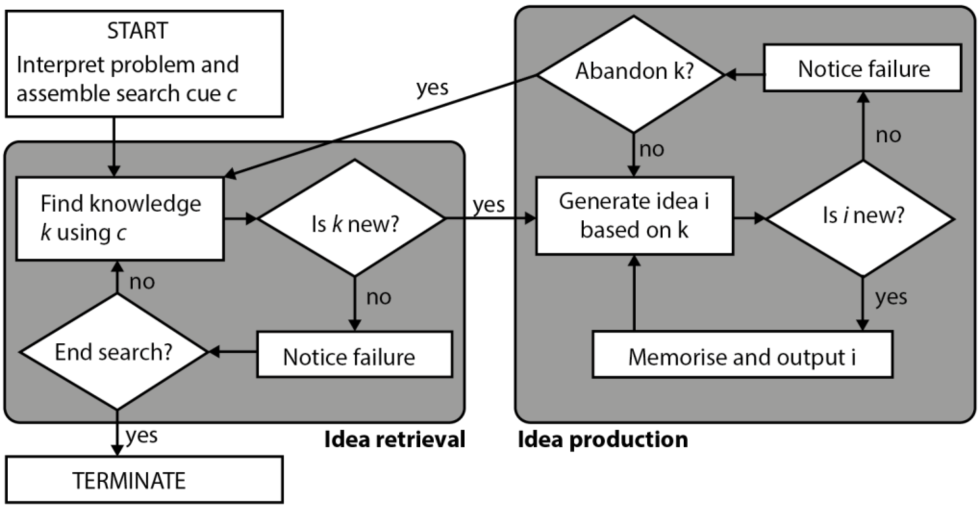


Figure 2-5 - Micro level of Model-L, redrawn from (Liikkanen, 2010)

The SIAM model was used as an initial hypothesis for design idea generation and elements of the model were validated or modified based on the effects observed in protocol studies and experiments. Two examples show how evidence was gained for the micro-level of Model-L in a design context (for a summary of the evidence, see Liikkanen (2010) p.70).

- Evidence of semantic and temporal clustering was found in design idea generation experiments, thereby supporting the existence of memory representations (Perttula and Liikkanen, 2006). Idea generation research has shown that consecutively generated ideas are often semantically clustered (come from the same category) and temporally clustered (it should take less time to generate ideas from the same category than different categories). SIAM explains this effect using the notion of memory representations[3]. In the memory retrieval phase, images are activated and one image can be used to produce different ideas based on that image, leading to semantically and temporally clustered ideas. (Perttula and Liikkanen, 2006) found evidence of semantic and temporal clustering in design, providing support for this aspect of Model-L.
- Unlike SIAM, which assumes that only one retrieved representation can be active at a time, in Model-L multiple representations can be active simultaneously,

[3] Originally called 'images' (Nijstad and Sroebe, 2006)

thereby accommodating combination. This adaptation (multiple representations) was based on verbal protocol in which a designer combined two representations (Liikkanen and Perttula, 2010). This suggests that individuals can maintain multiple simultaneous representations, in contradiction with the single-representation assumption of SIAM.

Although Model-L can accommodate combination processing and has received some empirical validation, the model does not provide an explanatory mechanism for the effects of combination, such as the effects of similarity on creativity or combination type. This may be contrasted with the Systematised Theory of Concept Generation (Taura and Nagai, 2013a, 2013b) that explicitly accounts for concepts similarity, dissimilarity and combination types at its core but has little in the way of empirical support for the nature of the underlying processes. This highlights the need for more empirical support for the systematised theory of concept generation, an extension of the micro-level of Model-L, or new models that build on the strengths of both.

2.1.3 Interim summary – design

In creative output production, combination involves the creation of a new design concept based on two more previously externalised entities. Combination can be described with many variables, as summarised in Figure 2-1.

The majority of knowledge about designer combination processes is at the computational level of description. 'Combination' as defined in protocol studies and their associated cognitive models is a relatively homogeneous construct, typically defined in the form of an input-output activity or transformations between two entities. The research on the phenomena of combination is relatively immature but some phenomena are beginning to emerge, focusing on the effects of base-concept similarity or distance on a variety of outcome measures including creative performance and combination 'type'.

There appears to be a disconnect between the computational and algorithmic levels of knowledge in the context of combination cognitive processes. There are algorithmic-level models of design ideation, but they do not explain the phenomena of combination. The most well-developed model, Model-L, does not account for any of the aforementioned relationships from similarity or relatedness. Conversely, research on the phenomena of combination rarely extends to consider algorithmic-level processes. In some protocol studies, authors decompose a combination process into subordinate processes (i.e., the processes involved in combination) that may be construed as an algorithmic-level representation, but they offer little in the way of explanatory mechanisms for computational level phenomena and there is no empirical evidence for any of these processes.

2.2 Combination cognitive processes in theories and models of creativity

Creativity in this thesis is defined as the ability to bring into being ideas, concepts products or outcomes that did not exist before. In research on creativity across multiple domains (including music, art, problem-solving, and design) a more specific definition of creativity is the production of outcomes that are novel and have some appropriate utility for the domain. To explore the role of combination in creativity, a search was conducted through the literature on theories and models of creativity that do not explicitly encompass design creativity (Kozbelt *et al.*, 2010; Lubart, 2001; Reiter-palmon *et al.*, 2015; Sowden *et al.*, 2015). Table 2-5 lists the models that were identified and which include some kind of combinational process.

The Geneplore (Finke *et al.*, 1992) and Creative Problem Solving (CPS) models (Mumford *et al.*, 1991, 2012) were reviewed in more detail as they were the only two that include algorithmic-level explanations for combinational cognitive processes. The Geneplore model includes two reciprocating phases of generation and exploration; the generative phase includes a cognitive process termed 'mental synthesis' that involves the combination or assembly of parts into wholes. The CPS model involves a 'conceptual combination' process associated with the combination and reorganisation of knowledge.

Table 2-5 - Creativity models

Model name (source)	**Structure of model**	**C.P**
(Guilford, 1968)	Divergent and convergent thinking	
BVSR / Constrained stochastic creativity (Campbell, 1960; Simonton, 1999)	Dual processes of (i) 'blind' generation of ideas via a stochastic process and (ii) selective retention of the best fitting ideas	*
Creative process model (CPS) (Mumford *et al.*, 1991, 2012)	Linear with feedback loops.	*
Geneplore (Finke *et al.*, 1992)	Two reciprocating phases of generation and exploration	*
(Runco and Chand, 1995)	A two-tier model with three sets of processes.	*
(Basadur, 1997)	Idea-evaluation cycles occur at each of three major stages: problem finding, problem-solving and solution implementation	
Dual-state model (Howard-Jones, 2002)	Associative and analytical thinking	
Honing theory (Gabora, 2004)	Flattened associated hierarchies during idea generation, refinement of an idea occurs "through interaction between the current conception of the idea and the individual's internal model of the world, or "worldview".	
Dual-pathway to creativity Nijstad (2010)	Creativity can arrive through two pathways: flexibility and persistence	

'C.P.' = models that describe combination processes

Both the Geneplore and CPS models involve a combinatorial cognitive process but each model focuses on different aspects of creativity. The CPS model (Mumford *et al.*, 1991, 2012) [4] was intended to "specify the critical processes typically involved in incidents of creative thought" (Mumford *et al.*, 2012, p.31). Processes are defined as "the operations, rules, and procedures guiding the application of knowledge in problem-solving" (Mobley *et al.*, 1992) (p.127). The cognitive processes involved in creative problem solving are

[4] The model has changed over time, originally including a 'combination and reorganisation' process as the only means of creative output production, but later making a distinction between 'conceptual combination' and 'idea generation' (Mumford *et al.*, 2012).

represented in a linear structure to show that the problem solving process has a start and end point, but they are also linked by feedback loops to show that people apply the processes selectively and cycle backwards and forwards whenever necessary.

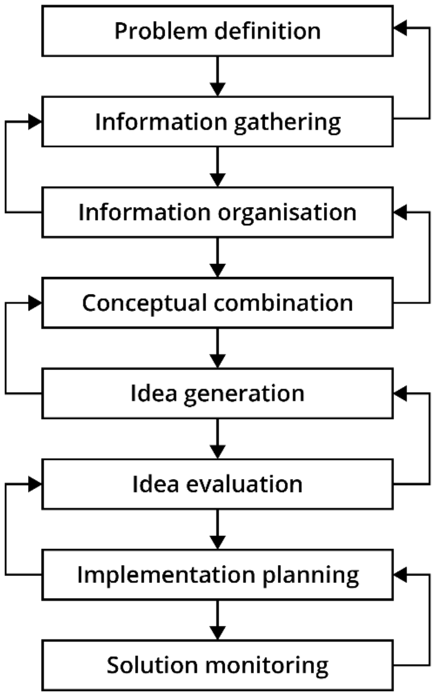


Figure 2-6 – The 2012 version of the creative process model (Mumford *et al.*, 2012).

The Geneplore model represents creative thought as two reciprocating phases of generative and exploratory thought bound by constraints. In the generative phase, individuals produce preinventive structures. These are initial, partially formed ideas or rudimentary structures that are subsequently interpreted as meaningful concepts and exploited in the exploration phase. These preinventive structures may be geometric forms, mental blends, category exemplars or more complex mental models (Finke *et al.*, 1992). As a model of creative thought generally rather than creative problem-solving, it does not address the problem definition and information gathering processes represented in the CPS model.

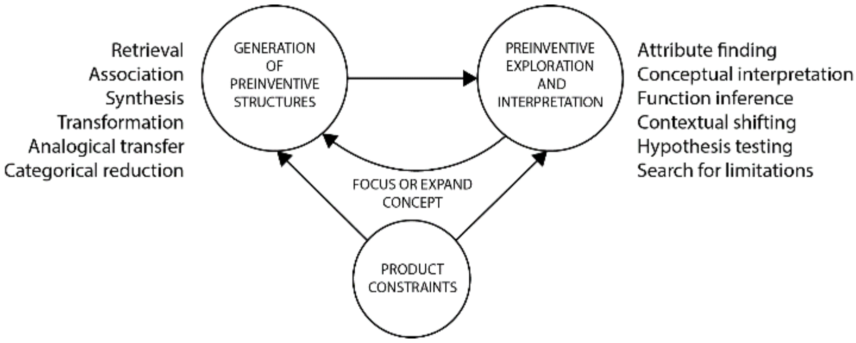


Figure 2-7 – The Geneplore model (Finke *et al.*, 1992) with generative and exploratory processes annotated next to the respective processing phase.

2.2.1 Computational level knowledge

The combination and reorganisation and mental synthesis processes both represent a person's capacity to create new entities through combination. The goal of creative combination is generally to create new outputs that are novel and useful. A variety of stimuli can be combined to produce creative outcomes. Studies of combination and reorganisation have used relatively simple category concepts (Baughman and Mumford, 1995; Mobley *et al.*, 1992) and more complex stimuli such as the combination of learning techniques to create new teaching methods (Scott *et al.*, 2005). According to the Geneplore model, people can mentally synthesise a variety of stimuli including visual patterns, object forms, mental blends, category exemplars, mental models and verbal combinations (Finke *et al.*, 1992).

Both processes are studied to identify factors that improve creativity and the outputs of the combination processes are typically measured in terms of e.g., quality and novelty. Research associated with the Geneplore model investigates how creativity can be improved by separating generative and exploratory processing and through the effective application of constraints. A variety of experimental tasks have been used to study different aspects of mental synthesis for different reasons (see Finke et al. (1992) for a review). In the context of creativity, a frequently used task is the figural combination task (Finke and Slayton, 1988). The task involves presenting stimuli such as geometric forms (Figure 2-8, left) to the participants and asking them to close their eyes and imagine combining them into a new recognisable form. These forms are then interpreted and developed into meaningful objects (Figure 2-8, right). Variants on the task manipulated variables such as the stimuli used, and the constraints placed on the task.

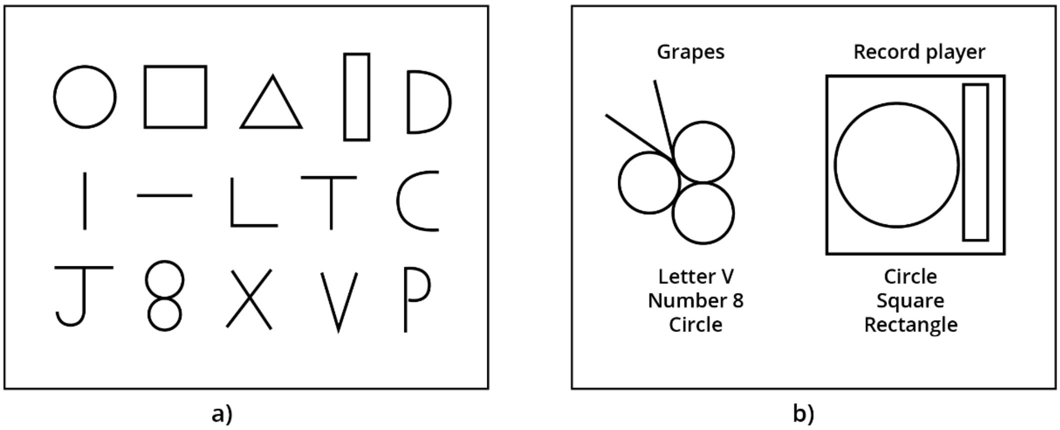


Figure 2-8 – an example of the 2D geometric parts (a), and outputs (b) of the figural combination task (Finke *et al.*, 1992; Finke and Slayton, 1988)

The figural combination task has been used to study e.g., spatial working memory (Pearson *et al.*, 1999) and the role of mental imagery and sketching in mental synthesis (Kokotovich and Purcell, 2000; Verstijnen *et al.*, 1998a). It has been proposed that there are two meaningfully different sub-types of mental synthesis (Verstijnen *et al.*, 1998a) (see also Verstijnen et al. 2000; I M Verstijnen, van Leeuwen, Goldschmidt, Hamel & J. M. Hennessey 1998). These are combination (overlaying and arranging parts to create new forms) and restructuring (manipulation of the individual parts by e.g. stretching or skewing). Two studies have also compared participants with and without design expertise (Kokotovich and Purcell, 2000; Verstijnen *et al.*, 1998a).

A key finding from early research associated with the CPS model is that originality and quality were influenced by the relatedness of stimulus categories. Participants were presented with exemplars from two or three taxonomic categories and are asked to create a new category that could account for all of the given exemplars (Figure 2-9) (Mobley *et al.*, 1992). More diverse categories contributed to an increase in the originality of outcomes, but the presentation of related categories improved quality. This finding prompted multiple studies to investigate the processes involved in combination and reorganisation which are discussed in the next section.

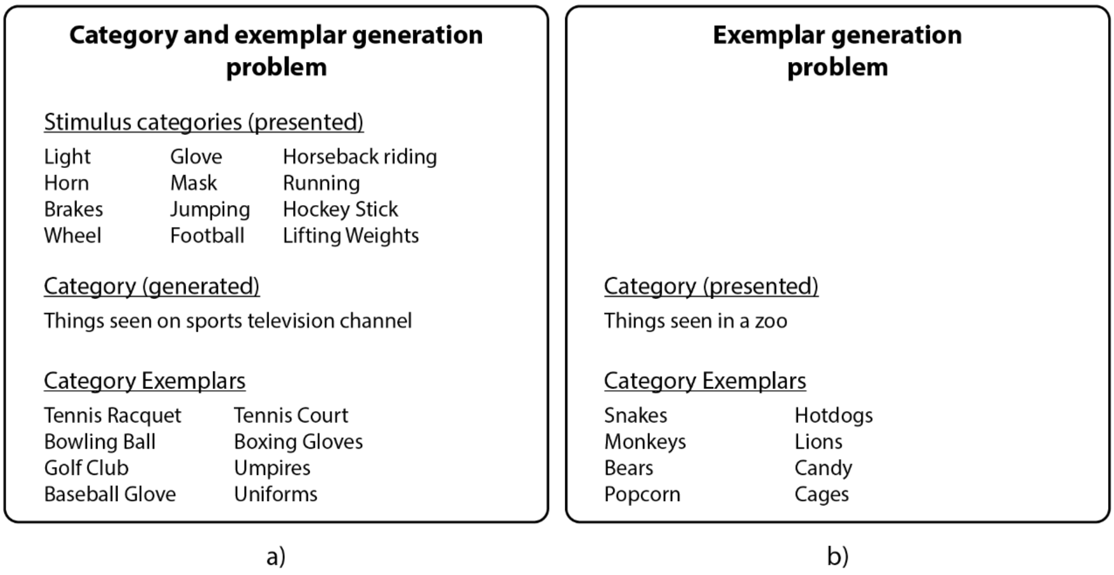


Figure 2-9 – An example of the category and exemplar generation task from Mobley *et al.* (1992)

2.2.2 Algorithmic level knowledge

An algorithmic level model of a creative combination process should explain how individuals combine entities and should account for effects identified in experiments. The majority of research at this level is associated with the CPS model.

The Geneplore model does not specify how mental synthesis occurs. Later research provided evidence for the role of mental imagery in synthesis and the distinction between combining and restructuring (Verstijnen *et al.*, 1998b, 1998a, 2000). They hypothesised that combining and restructuring would impose different loads on mental imagery and that sketching might differentially aid one process over the other. Results from a part-whole detection task and a modified version of the figural combination paradigm show that while *combining* parts is relatively easy and can be done in mental imagery, *restructuring* is more difficult and can be improved by experts using sketching. This differential effect of sketching is taken as evidence for two cognitive processes.

The combination and reorganisation process is said to occur via two sets of algorithms. A set of feature-search and mapping processes (similar to those involved in analogical reasoning) and case-based processes that draw from prior experiences and events rather than abstract features of entities.

Combination and reorganisation can occur through feature-search processes that operate on concrete or metaphorical features of the entities being combined. Initial evidence for the concrete feature-search processes came from Baughman and Mumford (1995). The authors proposed and tested a model of algorithmic-level processes that would explain the effects of category-relatedness on creativity (Mobley *et al.*, 1992). They proposed that to create new categories, people identify and extract features of the stimulus categories and use the common properties to define a new category. This was hypothesised to involve four processes:

(i) **Feature identification -** encoding and abstracting the properties of the categories,
(ii) **Feature mapping** – comparison or mapping of the properties that have been abstracted from each category,
(iii) **Category creation** – common or complementary properties identified through feature mapping are used in the definition of a new category,
(iv) **Feature elaboration** – individuals can now search for new concepts within the newly created category

Since 'diverse'[5] (low average similarity) categories will share few common properties it will be difficult to construct a new category that can accommodate the initial exemplars, thus leading to low-quality responses. However, searching for non-obvious relations in poorly related categories may lead to the production of more original outcomes. To test for these processes, instructions were given to experiment participants to induce the application of each process. It was found that (i) feature identification and mapping instructions (when performed together) and (ii) elaboration lead to the creation of more original category exemplars. This was taken as evidence for the existence of the processes. However, contrary to expectations (and the results of (Mobley *et al.*, 1992)), this only occurred when participants worked on similar categories.

The unexpected finding from the previous study suggested that there may be additional processes involved that were not captured by the four hypothesised operations. Mumford et al. (1997) hypothesised that individuals may not only extract features but use the abstracted metaphorical meaning of those features in the creation of a new category. Elicitation of feature mapping improved performance for individuals with closely related stimuli, but instructions designed to elicit the use of metaphorical meaning improved performance when individuals were working with diverse categories. Thus, when

[5] Operationalised as low average similarity from multiple independent raters

categories are sufficiently similar, combination may occur via feature-mapping mechanisms. However, more diverse categories require a search for more abstract relations to identify links between categories.

Conceptual combination may also proceed via case-based mechanisms (Scott *et al.*, 2005). Case-based reasoning is described as the abstraction and use of: goals, actions, outcomes, causes of goal attainment, contingencies, resources and constraints to guide the situation at hand. Participants were presented with either analogical (feature-mapping) processes or a set of five processes thought to be involved in case-based combination. The case-based processes were as follows.

(i) Characteristics of the problem situation are reviewed to identify the goals, causal factors that might lead to goal attainment, contingencies, resources and restrictions.
(ii) Prior cases are reviewed and their strengths and weaknesses would be assessed with respect to the current problem
(iii) Causes, contingencies, resources and restrictions from prior cases are used to create a preliminary model solution.
(iv) The model solution is used to forecast action outcomes.
(v) The forecasted outcomes are used to create a revised solution structure.

In the experiment, participants were asked to develop a new experimental teaching method for a fictitious school by combining two to four existing learning techniques using either analogical or case-based mechanisms. If analogical and case-based approaches to combination were cognitively distinct, it was expected that creative performance would be improved when stimuli were given in a form that was compatible with those specific processes. It was found that when concepts were presented in terms of their constituent features, the effective application of analogical (feature-mapping) mechanisms led to improved solution quality, originality, and elegance. When concepts were presented in a case format (goals, content and outcomes), effective application of the case-based heuristics led to improvement in the same measures. This showed that depending on the kind of knowledge being combined, creative combinations might be achieved using feature-mapping or case-based processing.

2.2.3 Interim summary – creativity

In the context of creativity, two models were found that give a particular focus to combination processes. The Geneplore model (Finke *et al.*, 1992) focuses on the role of mental imagery in creative thought and how creativity can be improved by separating generative and exploratory processes and placing constraints on the creative process. The Creative Process Model (Mumford *et al.*, 1991, 2012) focuses on how one must recombine and restructure one's existing knowledge to produce creative solutions to ill-defined problems.

At the computational level, both mental synthesis and combination and reorganisation represent an individual's general capacity to create new entities by combining stimuli. Research associated with the Geneplore model focused on how adding constraints can improve creativity, and research associated with the CPS investigated the role of relatedness and process execution on creative outcomes.

At the algorithmic level, combination may involve mental imagery (Verstijnen *et al.*, 1998a), feature-mapping processes (similar to those involved in analogical reasoning) using the concrete or abstract features of stimuli, or case-based processes in which multiple stored cases are combined to create model solutions to the problem at hand. Across all creativity research, knowledge is created using the hypothetico-deductive method to propose explanations for effects which are then tested using controlled experiments.

2.3 Linguistic conceptual combination

The third of the three combinational phenomena is termed 'linguistic conceptual combination'[6]. Conceptual combination in this context refers to a process in which a pair of 'base concepts' are interpreted to produce a "novel entity that is more than the simple sum of its component parts" (Ward *et al.*, 1997, p. 6). For example, when presented with the base concepts 'saw scissors' a person may interpret them to mean a pair of scissors where one or both blades have a serrated edge that can be used for sawing (Wilkenfeld and Ward, 2001). The types of combination presented in this review section are noun-noun artefact combinations as these most closely reflect the kind of artefact combination that occurs in design. There can also be adjective-noun combinations and the concepts can

[6] Commonly referred to simply as 'conceptual combination' by the authors cited in this section

also be, for example, biological entities (e.g. maroon carrot). In Section 2.3.2, five models of conceptual combination are presented and summarised according to the computational and algorithmic levels of description.

Table 2-6 – Five models of conceptual combination, abridged from Lynott and Connell (2010)

Model	Author(s)	Conceptual knowledge (from (Lynott and Connell, 2010), p.11)	Output types	Processes
CARIN	Gagné, 2000, 2001; Gagné & Shoben, 1997, 2002	"Not specified, although includes distributional knowledge of relation frequency".	Relational	A process operates over a finite set of thematic relations.
Dual-Process	(Wisniewski, 1997a)	"Amodal schemata with slots and fillers (but see Storms and Wisniewski (2005)".	Property Hybrid Relational	Two parallel processes: • Property transfer • Relation linking
Interactive Property Attribution	(Estes and Glucksberg, 2000)	"Amodal schemata with slots and fillers".	Property Hybrid	Interpretations guided by interactions between head and modifier
Constraint model	(Costello and Keane, 2000, 2001)	"Not specified, but modelled as amodal schemata with slots and fillers".	Property Hybrid Relational	Constraint satisfaction process
ECCo	(Lynott and Connell, 2010)	"Linguistic distributional information and situated simulation of meshed affordances".	Destructive Non-destructive	Affordance meshing

2.3.1 Computational level conceptual combination

A consensus among all models of conceptual combination is that the process serves a communicative purpose and that there is "a correspondence between linguistic words and psychological concepts" (Ran and Duimering, 2009, p. 61). Combinations can be used to create new concept categories, convey information in a precise manner, or refer to specific entities within a particular context (Wisniewski, 1997b).

The communicative goals of conceptual combination inherently place constraints on the outputs, and these constraints are different between models. Wisniewski (1997b) proposes that the speaker and listener assume three constraints: (i) a combined concept refers to a new category of concept that differs somehow from the inputs, (ii) the source of the difference is derived from the modifier (in English this is the first of the two nouns, but differs in other languages), and (iii) the entity that the output refers to still has many commonalities with the head noun. It has been shown however that the modifier need not always alter the head (Lynott and Connell, 2010). In constraint theory (Costello and Keane, 2000), combinations are said to be the result of a process that seeks to satisfy three constraints: (1) Plausibility – that the given pair of words is intended to refer to something that the interpreter can understand and find plausible, (2) Diagnosticity – that the two base terms are the best terms to use to communicate the desired concept, and (3) Informativeness – that both base concepts are necessary, and the interpreted concept is thus more informative than the base concepts.

The inputs to conceptual combination are two 'concepts' (in this case object-object pairs), frequently represented by words but similar principles can also apply to e.g. pictorial representations of novel artefacts (e.g., Wisniewski and Middleton (2002)). The order of presentation of base concepts is important, as the listener may arrive at a different conclusion depending on the order in which the same two words are presented. The output of conceptual combination describes a new referent concept. Models of conceptual combination classify the outputs of the process into types but differ in what counts as a distinct type. The most diverse classifications describe three types of output (Costello and Keane, 2000; Wisniewski, 1997a) but other models assume two (Estes and Glucksberg, 2000; Lynott and Connell, 2010) or one (Gagné, 2001). Assumptions about what constitutes a distinct type of combination are important in conceptual combination research because they are often used as a dependent variable and are assumed to provide insight into the algorithmic-level processes that gave rise to them.

Lynott and Connell (2010) summarise the phenomena of conceptual combination. For example, similarity has been shown to influence both the degree of emergence and the type of combined concept. Relatively more similar concepts receive fewer emergent features than more dissimilar pairs (Wilkenfeld and Ward, 2001) and are more likely to be combined by transferring properties between two concepts (Wisniewski, 1996, 1997b) or by performing destructive meshes (Lynott and Connell, 2010), depending on which model one is using to define combination types. Context has been shown to affect how word pairs are interpreted (Maguire *et al.*, 2010) and how fast they are interpretated (Gerrig and Bortfeld, 1999). Lynott and Connell (2010) list other phenomena of conceptual combination, but they are not included here either because the evidence is mixed (compositionality) or because the relevance to design combination is not clear (relation frequency).

2.3.2 Algorithmic level

Algorithmic level models of conceptual combination are developed through iterative, hypothetico-deductive research. In this section the key elements of the models in Table 2-6 are summarised, the empirical evidence that supports the hypothesised processes is indicated where relevant and some disagreements between models are highlighted to illustrate the extent of consensus amongst models.

In the Dual-Process model, conceptual combination occurs via two parallel, competing cognitive processes that take advantage of taxonomic and thematic knowledge (Wisniewski, 1997b). Property and hybrid interpretations involve the transfer of features between concepts through a structural alignment process also involved in similarity judgements, analogy, and metaphor (Gentner and Markman, 1997). Relational combinations are said to be produced via a process of scenario creation that involves the binding of two concepts in a scenario via a thematic relation. Combination type is said to be determined by similarity, diagnosticity and thematic role plausibility (Wisniewski, 2000).

The interactive property attribution model (IPA) (Estes and Glucksberg, 2000) is an alternative to the Structural Alignment model of property attribution interpretations. It was predicated on the proposition that property interpretations occur when one concept has a *salient* feature that is *relevant* to the other and relational combinations occur in the absence of such complementary features. It was proposed that not only are salience and relevance important but that, in disagreement with the Dual-Process model, the similarity of a pair has no bearing on comprehension. Wisniewski refuted a number of the claims made by Estes and Glucksberg, but accepted the importance of feature salience and diagnosticity in the Dual-Process model and described the role of these factors in the Dual-Process model, suggesting that the likelihood of property interpretations is influenced by similarity, modifier diagnosticity and thematic implausibility (Wisniewski, 2000).

Costello and Keane (2000) proposed that the interpretation of noun-noun combinations is the result of the satisfaction of three constraints (plausibility, diagnosticity and informativeness) and three steps:

1. Creation of alternative interpretations

2. Analysis of each interpretation with regards to diagnosticity and plausibility
3. Evaluation of informativeness and determination of overall applicability

The model conflicts with the Dual-Process model regarding which properties are selected during property attribution and how property interpretations are produced. Costello and Keane (2001) claim that diagnosticity, not alignability influences the selection of properties for attribution. Wisniewski provides alternative explanations for the role of diagnosticity and proposes that it is compatible with an alignment view.

In the Embodied Conceptual Combination model (ECCo), conceptual combination involves the meshing of perceived affordances (Lynott and Connell, 2010). The model is based on an embodied view of cognition wherein concepts are represented by numerous sensory modalities, including perceptual, motor, and affective information. In ECCo, each input concept has affordances derived in part from past experiences and also created "on the fly" (Lynott and Connell, 2010, p.4). Meshing involves the coming together of mutually eligible affordances. In destructive combinations, affordances are meshed even if one of the concepts needs to be extensively modified or destroyed, whereas in non-destructive combinations both concepts remain relatively intact.

2.3.3 Interim summary – conceptual combination

Five models of conceptual combination have been reviewed, providing an overview of contemporary cognitive models that focus on a variety of aspects of conceptual combination and in some cases build upon previous models. These are the Competition Among Relations in Nominals (CARIN) model (Gagné, 2001; Gagné and Shoben, 1997), the Dual-Process model (Wisniewski, 1997b), the Interactive Property Attribution (IPA) model (Estes and Glucksberg, 2000), the Constraint Satisfaction model (Costello and Keane, 2000) and the Embodied Conceptual Combination (ECCo) model (Lynott and Connell, 2010).

At the computational level, conceptual combination typically involves the interpretation of two concepts represented as words. There is some consensus about the communicative purpose of conceptual combination, some of the constraints on the process and experimental effects.

As Wisniewski (1997a) notes, most of the research on conceptual combination focuses on the algorithms that transform the inputs into outputs. Knowledge is created through hypothetico-deductive research where hypotheses are proposed to explain effects and tested with controlled experiments. Despite numerous iterations over at least twelve models (for reviews see: Lynott and Connell (2010), Ran and Robert Duimering (2009)) there is still disagreement about how conceptual combination occurs. Estes (2003a) proposed that the Dual-Process model (Wisniewski, 1997a) represents a sound basis for a general model of conceptual combination but acknowledges that "some details of the model are in need of specification or revision" (p.315). However, new models have continued to be proposed and no unified, comprehensive model has been produced. A key feature of conceptual combination research is the classification of the outputs of combination into 'types' and the drawing of inferences about the processes that gave rise to those types depending on the characteristics of the types.

2.4 Cross-domain comparison: gaps in knowledge and the potential for knowledge transfer

The knowledge about combination cognitive processes from design, creativity and conceptual combination was compared to reveal the gaps in knowledge in design and opportunities for future research, thereby addressing Obj.1.

2.4.1 A comparison of combinational cognitive processes across three domains

Table 2-7 summarises the computational level knowledge, effects, and algorithmic level knowledge and research methods associated with combination cognitive processes in each domain.

2.4.1.1 Computational level knowledge and research methods

Common to the processes in all three domains is the creation of a new concept based on two or more inputs. Common across the three domains is that people can combine a variety of stimuli to create new outputs, including geometric stimuli and category concepts. In design and creativity, people also combine their prior solutions to problems to create new ones. This shows that cognitive processes of combination incorporate depictive visual information with descriptive semantic knowledge across both idea production and interpretation processes.

Beyond the commonalities, there are some key differences between the domains. In combination in design and creativity, whilst the degree of requirements and constraints placed on combination may vary, individuals are always attempting to produce new ideas in a form of divergent thinking. That is, they are trying to come up with new, alternative ideas that are novel or useful and may address functions or solve problems depending on the context. In conceptual combination, however, the listener has communicative goals and is attempting to arrive at a 'correct' interpretation. Whereas multiple different outputs may be beneficial in design and creativity, multiple construals would be undesirable in conceptual combination.

Phenomena and effects. Some common effects have only been studied in a single domain. For example, the finding from the mental synthesis experiments that creativity can be improved with the addition of constraints (Finke *et al.*, 1992) does not appear to have been extended to design combination or conceptual combination. Nor does the comparison of mental synthesis with and without sketching (Kokotovich and Purcell, 2000; Verstijnen *et al.*, 1998a).

Other phenomena have been observed across the domains. For example, 'far' combinations, i.e., those from base concepts that are relatively more dissimilar, unrelated or distant pairs of base concepts have been shown to lead to more novel outcomes in design (Chan and Schunn, 2015; Doboli *et al.*, 2014) and creativity (Baughman and Mumford, 1995; Mobley *et al.*, 1992). Related to novelty, in conceptual combination, more dissimilar concepts are associated with more emergent features, i.e., *novel* features not present in either base concept (Wilkenfeld and Ward, 2001).

Combination types have been investigated across all three domains. All models of conceptual combination that were reviewed (Table 2-6) define 1 or more types of combination. Multiple factors have been associated with determining output type, including similarity (Lynott and Connell, 2010; Wisniewski, 1997a), diagnosticity and thematic role plausibility (Wisniewski, 1997a, 2000) and the salience and relevance of concept properties (Estes and Glucksberg, 2000). Combination types from conceptual combination have informed the development of types of combinations in design (Doboli *et al.*, 2014; Nagai *et al.*, 2009), but it is not clear what factors determine combination type. In mental synthesis tasks, design and non-design students produce combining or restructuring responses (Verstijnen *et al.*, 1998a). Restructuring is aided by sketching, but combining is not.

There are unique characteristics of combination processing in design that are not found in other domains. Doboli et al. (2014) measure two kinds of combinations: property-based and relation-based combinations, analogous to the property-mapping and relation-linking interpretations from the Dual-Process model. They propose that there are relatively more relation-linking combinations than property-mapping combinations because the entities being combined have "orthogonal functionalities with few overlapping features" (p.99), or that designers may be influenced by relational combinations found in design methodologies.

It should be noted that there are inconsistencies regarding constructs and their operationalisations in the literature. For example, Doboli *et al.* (2014) claim to manipulate relatedness and similarity but they do not report any criteria or procedure for how the stimuli were created. Nagai *et al.* (2009) refer frequently to 'similarity' but they actually measure the number of commonalities and differences between concepts. This mirrors similar issues in the other domains, Mobley *et al.* (1992) claim to manipulate 'relatedness' to create 'diverse' knowledge structures but operationalise this using ratings of similarity.

Research methods. Research in design is carried out via quasi-experiments, protocol analysis or analysis of naturalistic datasets. Cognitive models are typically created, at least in part, from analyses of design protocol (Chan, 1990; Gero and Kannengiesser, 2004; Jin and Benami, 2010; Jin and Chusilp, 2006; Kim *et al.*, 2010; Kruger and Cross, 2006; Liikkanen and Perttula, 2010; Sim and Duffy, 2003; Stauffer and Ullman, 1991; Ullman *et al.*, 1988). Some models have used data from think-aloud and experimental designs (Jin and Benami, 2010) or a mixture of independent protocol analyses and experiments (Liikkanen and Perttula, 2010; Taura and Nagai, 2013b). Research in non-design domains is characterised by hypothetico-deductive research using experiments or quasi-experiments. Cognitive models and experiments are closely related; knowledge is created by identifying effects and proposing and testing models that explain those effects.

Table 2-7 – A comparison of combinational cognitive processes across three domains

Variable	Design	Creativity	Conceptual combination
Computational level			
Goals, requirements and constraints	Idea production (design concepts) Specific requirements vary depending on task instructions in a given experiment. Designers may derive functional requirements from problems or be given explicit requirements to satisfy. Combination may occur in the absence of requirements e.g.,(Nagai *et al.*, 2009). Constraints may be given to the designer or self-derived.	Idea production Specific requirements vary depending on task instructions in a given experiment. Creativity can be improved by separating generative thought from exploratory thought (consideration of constraints) but constraints lead to improved creativity	Interpretation Three constraints (Wisniewski, 1997b): (i) combined concept must refer to a new category that differs from the inputs, (ii) the source of a difference is derived from a modifier, and (iii) referent concept has many commonalities with the head noun.
Input concept type	Geometric forms (Kokotovich and Purcell, 2000); Category concepts (Nagai *et al.*, 2009); Prior solutions (Chan and Schunn, 2015)	Geometric forms (Finke *et al.*, 1992); category exemplars (Mobley *et al.*, 1992); learning techniques (Scott *et al.*, 2005)	Novel pictorial forms (Wisniewski and Middleton, 2002); category concepts (Wisniewski, 1996)
Outputs	Design concepts (Nagai *et al.*, 2009)	Meaningful objects (Finke *et al.*, 1992); category concepts (Mobley *et al.*, 1992); teaching methods (Scott *et al.*, 2005)	An interpretation of the communicator's intent in the form of a new concept
Phenomena and effects			
Effects on creativity	Mixed evidence for effects of combination distance on creativity (Section 2.1.1.4).	More diverse categories contribute to an increase in the originality of outcomes, but the presentation of related categories improves quality (Mobley *et al.*, 1992).	---
	---	Effective application of constraints can lead to improved creativity (Finke *et al.*, 1992).	---

	Positive correlation between creativity and (i) combining score and (ii) restructuring score in the 3d figural combination task (two groups: industrial design engineering students and novices) (Verstijnen *et al.*, 1998a)		---
Effects on combination type	Sketching aids restructuring but not combination (Verstijnen *et al.*, 1998a) (Engineering design students)	Sketching aids restructuring but not combination (Verstijnen *et al.*, 1998a) (non-design students)	---
	---	---	Relationship between similarity and combination type (Wisniewski, 1996)
Effects on emergence	Non-linear relationship between combination originality and emergence (Nagai *et al.*, 2009)	---	Relatively more similar concepts have fewer emergent features and vice versa (Wilkenfeld and Ward, 2001)
		Algorithmic level models	
	Combination as a limited form of idea generation carried out via the Micro level of Model-L (Liikkanen, 2010; Liikkanen and Perttula, 2010)	Feature search and mapping (Baughman and Mumford, 1995; Mobley *et al.*, 1992); Case-based mechanisms (Scott *et al.*, 2005); mental imagery processing (Verstijnen *et al.*, 1998a)	Affordance meshing (Lynott and Connell, 2010); Causal reasoning or mental simulation (Wilkenfeld and Ward, 2001); Constraint satisfaction (Costello and Keane, 2000); Combination via a finite set of thematic relations (CARIN) (Gagné and Shoben, 1997); Dual processes of Comparison via structural alignment and Scenario creation via slot-filling (Wisniewski, 1997b); Salience and relevance interactions between head and modifier (Estes and Glucksberg, 2000).
		Research methods	
	• (Quasi-) experiments • Protocol analysis • Retrospective analysis of naturalistic datasets	• Experiments	• (Quasi-) experiments

References in the table are examples and not comprehensive lists of all relevant sources.

2.4.1.2 Algorithmic level knowledge

Representations of cognitive processes at the algorithmic level explain *how* input-concepts are combined to create a new output. Examples of algorithmic level representations are found across all three domains. Table 2-7 shows which algorithmic level processes were identified in each domain and Table 2-8 provides summaries for each process.

In design, algorithmic level representations are found in subdivisions of protocol codes (Jin and Chusilp, 2006; Kruger, 1999; Kruger and Cross, 2006), subdivided processing steps for the three forms of combination in the systematised theory of concept generation (Taura and Nagai, 2013a, 2013b) and the micro level of Model-L (Liikkanen, 2010; Liikkanen and Perttula, 2010). In models of creativity and conceptual combination, algorithmic level processes are found in the operations (strategies) underlying combination and reorganisation (Mobley *et al.*, 1992) and the analogical and case-based strategies that can be involved in combination (Scott *et al.*, 2005). In conceptual combination, all five models listed in Table 2-6 provide different algorithmic-level accounts of conceptual combination cognitive processing. As noted by Wisniewski (1997b, p.167) and as can be seen in the review by (Ran and Duimering, 2009), the algorithmic level is the level at which “most psychological approaches to understanding conceptual combination have been formulated”.

Research into design combination has borrowed algorithmic-level concepts from other models. Wisniewski's (1997b) Dual-Process model was used as a partial basis for four experimental studies of design combination (Doboli *et al.*, 2014; Jang, 2014; Jang *et al.*, 2019; Nagai *et al.*, 2009) and in one case (Doboli *et al.*, 2014) this was combined with research associated with the Creative Process Model (Mumford *et al.*, 2012). A common feature of this interdisciplinary knowledge transfer is the piecemeal adoption of elements of existing models. Authors use these existing models to formulate research questions or hypotheses, but they avoid making any falsifiable predictions about the existence of those processes in design.

Table 2-8 – Algorithmic level models of combination cognitive processes

Process	Description
Abstraction	Abstraction is thought to be involved in the category-generation task (Baughman and Mumford, 1995). When presented with diverse (dissimilar) stimuli in category combination, individuals will first attempt to extract features from the constituent categories via feature mapping process and if this fails they will turn to using abstract, metaphorical features of their concepts.
Affordance meshing	In the ECCo model of conceptual combination, (Lynott and Connell, 2010) propose that interpretation occurs via a process of meshing the perceived affordances of two concepts. This process is based on an embodied view of cognition.
Causal reasoning or mental simulation	Emergent features in conceptual combination may be attributable to causal reasoning processes that are engaged to explain incongruences that arise when two concepts are combined (Wilkenfeld and Ward, 2001). Alternatively, mental simulations may be used to address conflicts that arise in combination.
Case-based mechanisms	Case-based reasoning is the use of specific knowledge of past problem situations (cases) for use in solving a new problem (Aamodt and Plaza, 1994). It has been shown that in tasks designed to elicit combination and reorganisation, an alternative to the use of analogical mechanisms is the use of case-based reasoning mechanisms (Scott et al., 2005). This involves the "identification of outcomes, and influences on these outcomes, followed by forecasting and revision" (p.94)
Constraint satisfaction	According to the model of constraint satisfaction (Costello and Keane, 2000), interpretation in conceptual combination is determined by a process that attempts to satisfy three constraints: informativeness, diagnosticity and plausibility
Comparison via structural alignment	Property and hybrid interpretations in conceptual combination (Wisniewski, 1997a) are proposed to occur via the alignment and transfer of properties between concepts.
Relational element theory	Combination and reorganisation (Baughman and Mumford, 1995; Mobley et al., 1992) (based on relation element theory, Herrmann and Chaffin (1986)) similarly involves multiple comparisons to search for common elements amongst multiple concepts that can be used in the creation of a category that encompasses the given stimuli. Relations are constructed on-the-fly when two categories are brought together.
Scenario creation	Scenario creation processes involve 'combination' by placing base concepts into complementary roles in a relation (Wisniewski, 1997a). This process is derived from models of conceptual combination.

2.4.2 Gaps in knowledge

Current knowledge about combination cognitive processes in design is predominantly at the computational level. A comprehensive understanding of the effects of combination would require a systematic, experimental exploration of all possible permutations listed in the computational level framework (Figure 2-1). Any configuration of these variables not yet studied could present an opportunity for future research, and there may be more variables (see Vasconcelos and Crilly (2016)) that influence combination that have yet to be recorded. However, there are three salient gaps in knowledge when comparing the knowledge about design combination with other domains.

- There is a lack of knowledge about the combination of *design concepts*. Most outcome-based combination experiments in the design literature investigate the combination of category concepts (i.e., inspirational stimuli); only one outcome-based study (Chan and Schunn, 2015) investigates the combination of design concepts (solutions) as is prevalent in protocol studies.
- There is a lack of knowledge about how design experience influences combination processing (i.e., designers vs non-designers). For example, Doboli et al. (2014) propose that the relatively larger quantity of relational combinations may occur designers may be influenced by relational combinations found in design methodologies.
- There is a lack of knowledge about the influence of constraints and requirements on the outcomes of combination such as creativity, emergence and combination type. In protocol studies, designers create ideas in response to requirements and to satisfy constraints, however experiments that elicit combinations have minimal requirements and no constraints.

At the algorithmic level, there is no empirically supported algorithmic level account of combination processes in design. This can be demonstrated from two perspectives

- None of the algorithmic level representations of combination cognitive processes have any empirical support. The model of information categorisation (Kim *et al.*, 2010) is purely theoretical, blending two existing cognitive models but not evaluating them in a design context. In the research by Taura and Nagai (Nagai *et al.*, 2009; Taura and Nagai, 2013a, 2013b), processes can be distinguished by their *outputs*, but the nature of the associated process is purely hypothetical. Doboli *et al.* (2014) use the Dual-Process model of conceptual combination to generate hypotheses, but does not test any predictions of the model associated with algorithmic level processing.
- Empirically supported algorithmic level processes do exist, but they have not been validated for combination. The micro-level of Model-L (Liikkanen, 2010; Liikkanen and Perttula, 2010), is the most granular account of designer cognition supported by empirical evidence and describes idea production in a similar way to algorithmic level processes in conceptual combination and creativity research. However, it does not provide any explanations for the combination effects found in

the literature (Section 2.1.1.4) and has not been validated in the context of a combination experiment.

In summary, the current state of knowledge about combination cognitive processes is predominantly descriptive with some evidence of effects at the computational level, but with no empirically-supported algorithmic-level explanations for how designers turn inputs into outputs during combination.

2.4.3 Opportunities

The knowledge of cognitive processes and methods present in non-design domains presents opportunities for new research in design. By comparing the three domains, it is apparent that in the domains of conceptual combination and creativity research there have been relatively more proposals for algorithmic-level models of combination cognitive processes. This is especially true in the research about combination and reorganisation (Baughman and Mumford, 1995; Mobley *et al.*, 1992; Scott *et al.*, 2005) and the five models of conceptual combination that were reviewed (Table 2-6). There is an opportunity to use similar hypothetico-deductive methodologies and (quasi-)experimental research methods to advance the state of knowledge in design. Cognitive models can also be used as the starting point for the development of new models of design cognition that can be tested experimentally. As noted in Chapter 1, in this thesis this was done via a process of analogical abduction.

The challenge with this opportunity is how to select relevant models and methods as starting points for research in design. This is due to the wide range of algorithms that have been proposed as explanations for combination processes in other domains, the ongoing disagreement between models of conceptual combination (Ran and Duimering, 2009), and the relative lack of knowledge about designer combination processes.

The research from other domains offers two broad paths for future research. A research approach modelled after Mumford and colleagues would involve creativity assessment and the testing of the effectiveness of processing instructions. A research approach modelled after conceptual combination would explore combination types and their determinants. Ideally, an algorithmic-level cognitive model should be able to explain effects on both creativity and combination type and account for any unique characteristics of combination in *design* specifically. Existing descriptive research in design shows provides a starting point for both approaches. Common effects of combination distance on novelty (Section 2.4.1) would provide a starting point for mirroring the studies conducted by Baughman and Mumford (1995). It has also been shown that designers create multiple kinds of combination that align with the types of combinations described by the Dual-Process model of conceptual combination (Wisniewski, 1997a). The latter approach was adopted in this thesis, as discussed in the next section.

2.5 The case for a Dual-Process model of design concept combination

In this thesis, the Dual-Process model of conceptual combination (Wisniewski, 1997a) was used as an analogical target for proposing new cognitive models of designer combination cognitive processes. This section provides the rationale for this decision.

The work of Nagai *et al.* (2009) provided an initial indication that there may be some alignment between design combination and linguistic conceptual combination. It has been shown that when designers combine category concepts (such as guitar, ship or desk) to create design concepts, the resulting design concepts can be classified into three types that correspond to the outputs of the Dual-Process model of conceptual combination (Nagai *et al.*, 2009; Taura *et al.*, 2007; Taura and Nagai, 2013a). Although cognitive processes cannot be inferred solely from output types, the presence of these combination types provides a starting point for an empirical investigation into the processes that give rise to them.

There is also reason to propose a dual- rather than a single-process model of combination. First, there is converging evidence from diverse behavioural, neurological and computational research indicating that conceptual semantic knowledge is represented across two, dissociable featural (taxonomic[7]) and thematic systems (Mirman *et al.*, 2017). It is assumed that when designers combine category concepts or design concepts they are manipulating semantic concept knowledge. By extension, this semantic knowledge may also be processed by dual featural and thematic systems. Secondly, mental synthesis tasks

[7] (Mirman *et al.*, 2017) refer to taxonomically organised knowledge. Taxonomic semantic structure is based on the featural commonalities, with members of the same position in a taxonomic hierarchy sharing many common features with other members in the same position (Markman and Wisniewski, 1997). It has been proposed that design concepts are not taxonomically organisable (Taura, 2016c), but since taxonomic categorisation is based on the common features of concepts, the taxonomic organisation of semantic knowledge is taken to mean the feature-based organisation of semantic knowledge.

appear to elicit combination processing through two processes (Verstijnen *et al.*, 1998a, 1998b, 2000) that are similar to those in the Dual-Process model. The combination and *restructuring* of geometric forms is similar to comparison and construction, and combination *without modification* is similar to scenario creation. These two types of combination have been shown to impose different loads on mental imagery, suggesting that they not only produce distinct types of output but occur via distinct processes.

There is also evidence to support the proposition that design combination can occur via feature-based cognitive processing. Feature-mapping processes have been implicated in both creative combination (Baughman and Mumford, 1995) and conceptual combination (Wisniewski, 1997a). Designers are also seen to conduct feature-based combinations naturally. Daly *et al.* (2012) present an example of a designer combining the features of two concepts to create a new one (Figure 2-2). In this example, the magnifying glass feature of one concept has been added to the parabolic reflector of the second concept to create a new combined concept. Feature-mapping and transfer models offer plausible explanations for how this designer has transferred the magnifying glass feature between two concepts.

Finally, a Structural Alignment process such as the one involved in the Dual-Process model of conceptual combination provides a plausible basis for feature-based design combination. Designers are known to analogise in creative output production (Hay *et al.*, 2017a) which is thought to occur via the alignment and mapping of features between the base and target representations (Gentner and Markman, 1995, 1997; Gentner and Smith, 2012). It would be a parsimonious extension of this reasoning process if combination were simply a specialised form of analogy. Whereas analogical reasoning involves the mapping of old solutions to new problems, structural alignment in combination would operate on two 'solutions' to transfer features between them.

2.6 Summary

Chapter 2 has presented a literature review of combination cognitive processes in design, creativity and conceptual combination. The review was conducted to address Obj. 1 - to assess the state of knowledge about combination cognitive processes and identify research methods suitable for advancing that knowledge. To identify gaps in knowledge and opportunities for new research in design, the knowledge about *design combination* was compared with the knowledge about *non-design creative combination* and *linguistic conceptual combination.* Each domain was reviewed in terms of the computational level knowledge about combination cognitive processes at the computational level of representation (goals, inputs, outputs and effects) and the algorithmic-level (explanatory mechanisms that explain computational level phenomena). The resulting comparison highlighted the knowledge in non-design domains that has yet to be attained in design and the research methods that are used to gain that knowledge.

Two key findings were (i) the lack of knowledge about algorithmic-level combination cognitive processes in design, i.e., how designers can turn inputs into outputs through combination, and (ii) the opportunity for hypothetico-deductive, experimental research in design that builds on existing knowledge of cognitive processes in other domains. These two findings directly led to the proposal that the same processes involved in the Dual-Process model of conceptual combination (Wisniewski, 1997a) may also be involved in design combination, which in turn set the course for the research in the remainder of the thesis. The next chapter presents the Dual-Process model of conceptual combination in detail, alongside the related Dual-Process model of similarity judgements and background literature on human similarity judgements, both in general and in design.

3 An overview of human similarity judgements and the Dual-Process models of similarity and conceptual combination

In the previous chapter, the literature review of combinational cognitive processes revealed that there were no algorithmic-level models of combination cognitive processes in design. This led to the aim of modelling the cognitive processes of combination in conceptual PDE. It also highlighted the opportunity to use a hypothetico-deductive methodology and experimental research to create new knowledge in design (Section 2.4.3). To this end, a case was made for using the Dual-Process model of conceptual combination as the basis for a Dual-Process model of *design concept combination* (Section 2.5).

During the initial process of developing a research design to test the predictions of a Dual-Process model of design concept combination, the scope of the research was extended to include the additional objective of modelling the cognitive processes involved in *design concept similarity judgements* (Obj. 2). This was done for three reasons. First, it was proposed that similarity judgements and conceptual combination involve the same underlying cognitive processes, and knowledge about similarity judgements would provide insights into combination processing. Second, there was an apparent gap in knowledge about how designers make similarity judgements. Third, research into design concept combination would involve some research into similarity judgements anyway, and extending the research scope would be a relatively efficient use of resources.

This chapter provides a background review of knowledge about human similarity judgements generally and in design, highlighting the lack of knowledge about the cognitive processes involved in designer similarity judgements (Section 3.1). The next two sections present the two cognitive models that form the basis of new models of design cognition in Chapter 4. They are, the Dual-Process model of similarity judgements (Section 3.2) and the Dual-Process model of conceptual combination (Section 3.3). Finally, the aforementioned reasons for extending the research scope to include designer similarity judgements are discussed in more detail in Section 3.4.

3.1 Human similarity judgments

In accordance with the cognitive process framework introduced in Section 1.1.3, human similarity judgements are mental processes that act on mental representations to produce subjective impressions of similarity ("perceived similarity"). These mental representations are based on entities in the represented world and thus it may be assumed that there is a correspondence between perceived similarity and the 'real' properties of the entities in the external world. As will be seen in Chapters 8 and 9, similarity judgements from different people do display a central tendency, and so it may be assumed that there is some common human processes that give rise to impressions of similarity. That being said, people do not necessarily form the same mental representation of the same represented entity. A person's impression of similarity is highly context dependent and thus similarity judgements should be considered *relative* rather than *absolute* judgements. Human judgements of similarity may be distinguished from quantitative metrics of similarity, wherein items in the represented world are measured via some proxy for similarity and formulae are used to compute similarity in a (relatively more) objective manner.

3.1.1 Models of similarity judgements

Similarity judgements are fundamental aspects of human cognition, important for learning, knowledge, problem-solving, prediction and categorisation (Goldstone and Son, 2012). An understanding of how humans make similarity judgements facilitates the understanding of a multitude of facets of human cognition. Likewise, similarity judgements play a role in a range of cognitive processes used in design, and an understanding of how these judgements are made can aid in understanding these other processes. Following the review by (Goldstone and Son, 2012), models of similarity judgements may be classified as geometric, feature-set, feature-alignment, or transformational[8].

Geometric (or spatial) models treat entities as points in a metric space organised into dimensions. Within this space, perceived similarity is represented as an inverse function of the distance between those entities; the closer two things are in the space, the more similar they are.

In featural models, perceived similarity is a function of the number of common and different features shared by two concepts. Both feature-set and feature-alignment models (described as 'featural' models in this thesis, but terminology varies) assume that mental representations of concepts are composed of features. In feature-set models, concepts are represented as sets (lists) of features in the mind and perceived similarity is assumed to be a function of the common and different features of two concepts. In feature-alignment models, concepts are represented as structured representations, where similarity is not only a function of the common and different features but also of how those features are related. In all featural models, the more commonalities two things share, the more similar they should be judged to be.

In transformational models, the perceived similarity of two entities is determined by how complicated it is to transform one representation into another. This presumes that people conduct operations such as rotations, reflections, position swapping, mirroring or reversal to turn one concept into another.

[8] The terminology in this thesis differs from that used by (Goldstone and Son, 2012) to improve compatibility with terminology from conceptual combination research later in the thesis.

In this thesis, the focus is on featural models, particularly the Structural Alignment model (Markman and Gentner, 1993a) and its implicated role in conceptual combination (Wisniewski, 1997a).

3.1.2 The phenomena of similarity judgements

Experimental research has led to the identification of some phenomena of human similarity judgements. First, similarity is asymmetrical (Tversky, 1977), in that for two objects *a* and *b,* the similarity of *a* to *b* may be different than the similarity of *b* to *a.* This was one of the limitations of spatial models of similarity that Tversky's featural model of similarity (Section 3.2.1.1) aimed to overcome. Secondly, judgements of similarity ('how similar two things are') and difference ('how different two things are) are not always complimentary (i.e., the inverse of one another) (Golonka and Estes, 2009; Medin *et al.*, 1990; Simmons and Estes, 2008; Tversky, 1977). In featural models, both characteristics are attributed to the difference in the relative weight of commonalities and differences on the perception of similarity. Common features count more towards similarity than different features detract from it (Krumhansl, 1978; Markman and Gentner, 1993b, 1996; Tversky, 1977).

A second characteristic of human similarity judgements is that they are dynamic and context-dependent (Medin *et al.*, 1990; Tversky, 1977; Tversky and Gati, 1978). In feature-based models, the extent to which specific features influence perceived similarity may vary from person to person and one's perception of similarity can be influenced by the set of concepts under consideration (Tversky, 1977). For example, when asking people to compare the similarity of a set of cars, they may assume that the relevant frame of reference includes multiple aspects, such as form, function, cost, quality, size, performance etc., but the relative weights of these criteria may differ from person to person. A person's judgement of the relative similarity of two cars may differ when those two cars are viewed on the website of a single car manufacturer, versus a website containing many cars from all manufacturers (see the diagnosticity principle, (Tversky, 1977)).

A third characteristic is that there are individual differences in similarity judgements (Gentner and Brem, 1999; Golonka and Estes, 2009; Mirman and Graziano, 2013; Simmons and Estes, 2008). That is not to say that when people are, for example, asked to rate how similar two things are that they provide different ratings. Indeed, it should be expected that there is some natural variation or 'noise' in how people express their perceptions of similarity. Rather, individual differences refers to consistent behaviours shared by sub-groups in a population (see Section 3.2.3).

3.1.3 Similarity in design

Similarity judgements may be considered important in design for their involvement in, or common processing basis with, other cognitive processes. For example, similarity (or distance) has been conjectured to be a component of novelty judgements (Brown, 2016). As has been seen in Section 2.1.1.4, the creativity of the outputs of combination is related to the similarity of the inputs. This also extends to ideation via reasoning processes. The domain-similarity of target analogues has been shown to influence novelty (Chan *et al.*, 2011; Wilson *et al.*, 2010) and quality variation (Chan *et al.*, 2011) in analogical reasoning. (Chaudhari *et al.*, 2019) also propose that similarity is important for predicting artefact performance from the parameters of computational and physical prototypes.

Research on similarity in design may be considered in two groups of research, quantitative metrics or human similarity judgements. Some researchers have proposed similarity metrics that do not explicitly reflect human cognitive processing. McAdams & Wood (2002) developed a quantitative metric which computes the similarity of analogous products based on functional similarity and consumer needs. Fu et al. (2013) utilise latent semantic analysis to determine the semantic similarity of patent documents. Bao, Faas, & Yang (2016) propose a metric which measures the unique features of two concepts; the more unique features between a pair, the lower the similarity.

Other research approaches attempt to understand designer cognition, or use human judgements to evaluate quantitative metrics. Chaudhari *et al.* (2019) present a knowledge-based approach for similarity measurement in engineering design that accounts for the important role of scientific knowledge (e.g., knowledge of kinematics, material properties or heat transfer) typical in later stages of the design process. Nandy and Goucher-Lambert (2021) found that quantitative, feature-based measures of similarity align well with human similarity judgements. Gill *et al.* (2019) explored the dimensions along which existing products are judged to be similar to products being designed. Participants preferred to select products that were functionally similar from a predefined list of five dimensions (function, form, energy flow, material flow and motion). Ranawat and Hölttä-Otto (2009) show that colour, texture shape and form contribute to perceptions of product similarity.

The existing research that involves human similarity judgements has started to reveal the elements that contribute to similarity and metrics that may correspond to human judgements, but there have been no explicit attempts to create a cognitive model of designer similarity judgements. This highlighted the potential value of a cognitive model of designer similarity judgements, and contributed to the decision to extend the scope of the thesis research to include the objective of modelling designer similarity judgements of design concepts.

3.2 The Dual-Process model of similarity judgements

The model introduced here is the Dual-Process model[9] of similarity (Chen *et al.*, 2014; Estes, 2003a, 2003b; Estes *et al.*, 2011; Wisniewski, 1997a). According to this model, an individual's perception of similarity arises from one of two cognitive processes.

- A comparison process operates over the constituent features of a pair of concepts via a structural alignment algorithm. Similarity is a function of the commonalities alignable differences and nonalignable differences between the pair (Section 3.2.1).
- Thematic similarity is the perception of similarity that arises through a scenario creation process in which two concepts are perceived to be more similar because they are related by an extrinsic, complementary relation (Section 3.2.2).

The dual-process view of similarity can be viewed as an extension of earlier single-process featural models (Markman and Gentner, 1993b; Tversky, 1977) that only included a comparison process. The inclusion of a second process followed empirical evidence that thematic relations also influenced similarity judgements (Bassok and Medin, 1997; Wisniewski and Bassok, 1999). Wisniewski and Bassok (1999) drew on findings from conceptual combination to raise the need for a dual-process account, showing that contemporaneous research on conceptual combination influenced thinking on similarity judgements. However, there is ongoing debate regarding the extent to which thematic processing is a genuine form of similarity processing or whether it merely intrudes on comparison in certain circumstances (Gentner and Brem, 1999; Honke and Kurtz, 2019).

[9] Although the term 'model' is used here for consistency, unlike the Dual-Process Model of conceptual combination there is no singular work that can be attributed to proposing a model *per se.* Other authors may use the phrase 'dual-process view' of similarity.

3.2.1 Featural models: similarity via a comparison process

This section presents two featural models of similarity judgements, the contrast model (Tversky, 1977) and the Structural Alignment model (Markman and Gentner, 1993a). Both propose that similarity arises as a result of a comparison process that operates over the common and different features of a pair of concepts, but each differs in the specific algorithm and assumptions about mental concept representation. Only the latter comprises part of a Dual-Process model of similarity judgements, but both models are presented here because understanding the former facilitates understanding of the latter.

3.2.1.1 The Contrast Model

The contrast model assumes that concepts are represented in the mind (the representing world) as lists or sets of features. Similarity is a function of the matching and mismatching features at the intersection of these sets. This is visualised in Figure 3-1, common features (C) are those in the overlap, and different features (D) are those in only one circle. In this view of similarity, a DOG and HORSE would be similar because they have many features in common, e.g., both have a head, body, tail and legs. Similar models exist (Bush and Mosteller, 1951; Eisler and Ekman, 1959; Sjoberg, 1972) but all can be considered specialisations of the Equation 1 (Goldstone and Son, 2012).

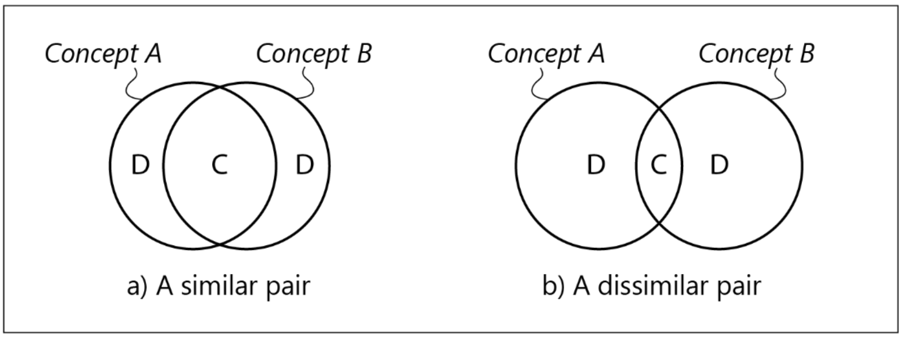


Figure 3-1 – Illustration of the relationship between commonalities (C) and differences (D) for similar and dissimilar pairs of concepts according to the Contrast model (Tversky, 1977).

According to the Contrast model, the similarity of a pair of items is computed by the equation:

Equation 1: $S(A,B) = \theta f(A \cap B) - \alpha f(A-B) - \beta f(B-A)$

That is, the similarity of A to B is a linear function of the common (A∩B) and distinctive ((A-B), (B-A)) features of the pair. Similarity is asymmetrical, in that the similarity of A to

B is not necessarily the same as the similarity of B to A. The model also makes correlational claims about pairs of variables:

- High similarity pairs should have many commonalities (matching features) and few differences (mismatching features).
- Low similarity pairs should have few commonalities and many differences.

3.2.1.2 The Structural Alignment Model

According to the Structural Alignment model of similarity (Gentner and Markman, 1994, 1997; Markman and Gentner, 1993b, 1993a), concepts are represented in the mind as structured representations and similarity arises from the alignment and comparison of these structured representations. Structured representations comprise:

- Entities – are arguments that refer to objects themselves (like nouns)
- Attributes – are arguments that describe objects (like adjectives)
- Relations – predicates that link two or more arguments (where arguments can be attributes, objects, or other relations). Relations that take other relations as arguments are called higher-order relations. Note that these relations are internal, rather than the external relations in thematic relations.
- Functions – values used when a statement cannot be true or false (e.g., a quantitative measure of size)

Figure 3-2 illustrates this concept using an example of geometric shapes adapted from Markman and Gentner (1996). The left concept comprises a triangle above a circle, both of which are beside a hexagon. The right concept comprises a circle above a triangle. The concepts are represented in terms of their entities and relations. This illustrates the 'relational structure' of each concept.

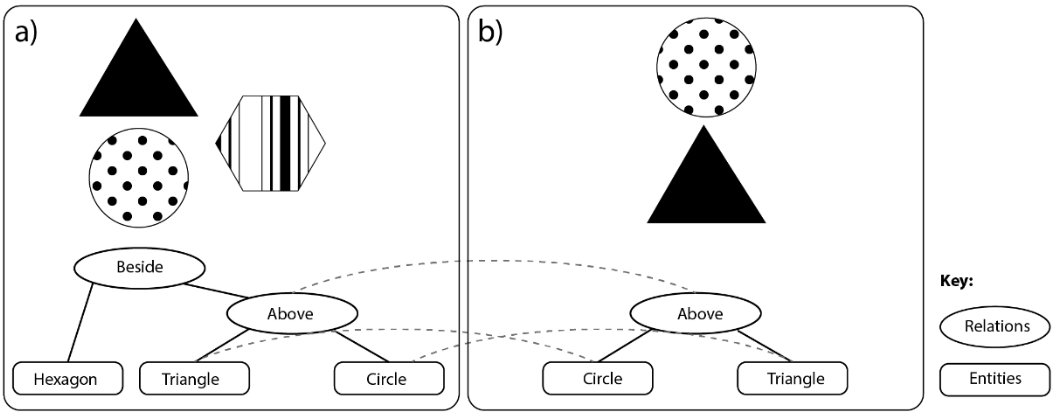


Figure 3-2 – Illustration of two aligned structured representations. Two concepts (a and b) comprise multiple geometric shapes. The concepts are represented as structured representations.

The process of structural alignment that gives rise to a perception of similarity involves the alignment of the relational structures of two concepts. The alignment process operates by attempting to achieve a *maximally systematic alignment* of the common relational structure of two concepts whilst adhering to two constraints; *parallel connectivity* and *one-to-one* mapping (Markman and Gentner, 1996). Parallel connectivity means that when matching relations are aligned, their arguments[10] are also aligned. One-to-one mapping means that a representational element of one concept can have no more than one matching representational element in the other concept when they are aligned. The preference for maximal systematicity means that the process will prefer to align matching relations over entities and higher-order relations over lower-order relations.

The alignment process enables the distinction between two kinds of differences: those that are related to the common structure (alignable differences), and differences that are independent of the common structure (nonalignable differences). In Figure 3-2, the two concepts are aligned at the 'above' relation, which facilitates the most systematic mapping between the two concepts. Parallel connectivity means that the arguments of the relation are also aligned (in this case the entities in the above relation). This means that the 'above' relation is a commonality and the circles and triangles become alignable differences. The hexagon and the 'beside' relation are non-alignable differences, because there is no common relational structure in (b) that it aligns with.

[10] The arguments of an aligned relation can be lower-order relations or objects

According to the Structural Alignment model, an individual's perception of similarity is a function of the degree of alignment of the two mental representations in terms of the commonalities, alignable differences and nonalignable differences. The difference between this model and all feature-set models is that similarity is a product of the comparison of not just the constituent features, but *also how they are related*. In this case, the similarity between DOG and HORSE is influenced by the relations between entities, e.g., how they are connected spatially and dynamically. This means it is not only important that both animals have legs, a head, a body, a tail and fur, but that the limbs and fur are all related to the body in the same way.

As the alignment process maximises the relational structure of two concepts, both commonalities and alignable differences become more salient, whereas nonalignable differences do not. This explains the previously established finding that commonalities count more towards similarity than differences, but also leads to the predictions that alignable differences should be more influential in the perception of similarity than nonalignable differences and should also be easier to list (and thus more numerous). The predictions made by the model are as follows (Markman and Gentner, 1993b, 1996).

1. Perceived similarity can be predicted as a function of commonalities, alignable differences and nonalignable differences, and alignable differences should be a greater influence on similarity than nonalignable differences.
2. Pairs with increasing perceived similarity should also have more commonalities
3. Pairs of concepts with many commonalities should also have many alignable differences.
4. Alignable differences should be easier to identify than nonalignable differences.
5. Commonalities should be conceptually related to the alignable differences.

As per the Contrast model, perceived similarity can still be predicted by a function of the number of commonalities and differences that can be listed for that pair (in this case, two kinds of differences) and relatively more similar concepts should have an increasing number of commonalities. Where the two views of similarity differ is in the relationships between perceived similarity and the number of differences in a pair of concepts.

The second and third predictions (above) lead to an apparent contradiction; concepts that are perceived as being highly similar should have many commonalities *and* many alignable differences, even though alignable differences detract from similarity. How then does similarity increase as a function of commonalities and decrease as a function of differences? Figure 3-3 is a novel illustration representing the influence of commonalities, alignable differences and nonalignable differences on perceived similarity that takes the relative influence or 'weight' of each variable into account, represented by the size of the circles. Commonalities count more towards similarity than differences detract from similarity (Krumhansl, 1978; Markman and Gentner, 1993b, 1996; Tversky, 1977), and alignable differences count more against similarity than nonalignable differences (Markman and Gentner, 1993b, 1996). Thus, even though noticing commonalities leads to the noticing of alignable differences, rated similarity can still increase even when the number of listed total differences increases.

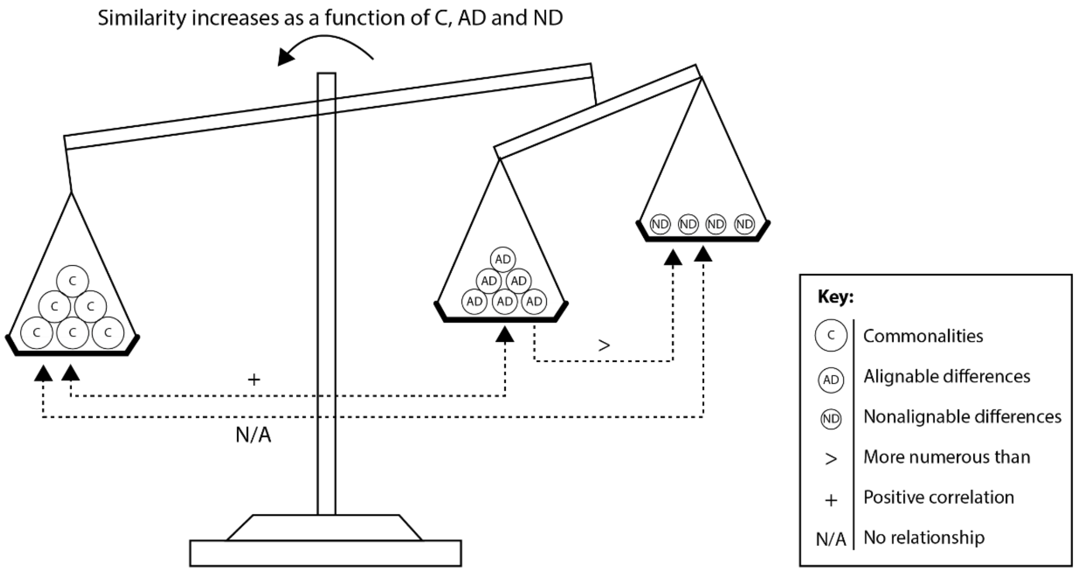


Figure 3-3 – An illustration to show the relative expected contribution of commonalities, alignable differences and nonalignable differences to rated similarity. Commonalities influence similarity more than ADs, and ADs count more than NDs.

3.2.2 Thematic similarity

The second process in the Dual-Process model of similarity can be termed 'thematic similarity'. With thematic similarity, people perceive two concepts as being more similar because they are thematically related.

There are different definitions of thematic relations in the literature and different views on how thematic similarity might manifest. The most permissive definition is the spatial or temporal co-occurrence of two concepts in some situation or event. For example, "BOWLING PIN and ARCADE" could co-occur in the same setting (Honke and Kurtz, 2019) or be members of the same thematic category (things found in a bowling centre). Such concepts may be judged as being more similar just by co-occurrence in an external relational structure (Jones and Love, 2007).

A more restrictive view is that thematic relations must be external and *complementary* relations between two concepts (Estes *et al.*, 2011). Complementary means that the two concepts fulfil different roles in a relation. In the previous example, a bowling pin may be said to have no clear complementary relation beyond co-occurrence, but a bowling ball and a bowling pin can be bound by e.g., the 'hit' relation. In similarity judgements, a BALL and PIN share few common features but may be rated as being similar because of the thematic relation and their presence in the bowling scenario.

In empirical studies, participants are seen to justify their similarity ratings based on the existence of thematic relations (Bassok and Medin, 1997) and pairs of stimuli with pre-existing thematic relations are judged to be more similar than thematically unrelated concepts (Wisniewski and Bassok, 1999). For example, MILK and COFFEE can be rated as being more similar than MILK and LEMONADE because of the thematic relation between milk and coffee (milk is used in coffee).

Unlike the previous feature-based similarity models, there do not appear to be any algorithmic level models that make explicit predictions about how thematic perceptions of similarity occur. In some views, thematic relations are stored in knowledge schemata. An individual may have knowledge of typical scenarios that concepts appear in (Wisniewski, 1997a), or the potential roles that concepts can play may be stored as semantic properties of an object (Jones and Love, 2007). An alternative view is that thematic relations may also be derived from perceived affordances (Estes *et al.*, 2011). For example, while an individual may know that hammers can be used for hitting nails (HAMMER and NAIL are placed into the HIT relation), if given a sufficiently solid and heavy object they may also be able to determine that e.g., bricks or stones can be used for hitting nails. This view implies that individuals can create thematic relations 'on-the-fly'.

3.2.3 The interplay of the two processes

The tendency for individuals to display feature-based or thematic processing in their similarity judgements has been attributed to stimulus compatibility (Murphy, 2001; Wisniewski and Bassok, 1999) and individual differences in process preference (Mirman and Graziano, 2012; Simmons and Estes, 2008).

'Stimulus compatibility' (Wisniewski and Bassok, 1999) refers to properties of stimuli that make them inherently more or less compatible with comparison or scenario creation. Pairs of concepts that are highly 'alignable', meaning they can achieve a relatively high degree of alignment of relational structure, are compatible with the comparison process but tend to occupy the same role in a relation, making them incompatible with the scenario creation process. Concepts with low alignability are difficult to compare in a meaningful way but are easier to integrate into a scenario. Feature-based and thematic processes can make independent or cumulative contributions to similarity. Figure 3-4 shows an example (Estes *et al.*, 2011) extrapolated from Golonka and Estes (2009), and Wisniewski and Bassok (1999). Thematic relations can have a relatively large influence on perceived similarity in the absence of common features, but provide a relatively small cumulative increase when concepts already have many common features.

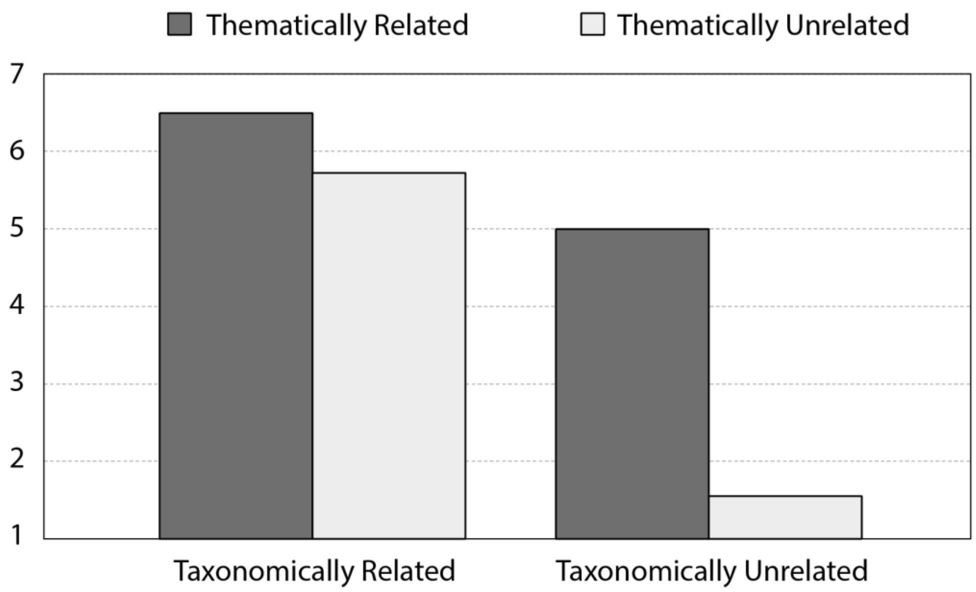


Figure 3-4 – Example of the 2x2 interaction of thematic relatedness and taxonomic category membership, redrawn from Estes *et al.* (2011).

Although the dual-process view of similarity implicitly assumes a generalizable influence of featural and thematic processing on perceived similarity, empirical research has identified individual differences in preferences for the two kinds of processing (Golonka and Estes, 2009; Simmons and Estes, 2008). Simmons and Estes (2008) found that the majority of people consistently favour either featural or thematic processing across forced-choice and similarity rating tasks. Subsequently, they found that when participants have a preference for thematic processing, their similarity judgements are more influenced by thematic relatedness than alignability, but when they have a preference for featural processing there is almost no effect of thematic relatedness on similarity. (Golonka and Estes, 2009).

A preference for thematic processing can be predicted by a participants' 'need for cognition' (the extent to which the participants tend to engage in and enjoy thinking (Cacioppo and Petty, 1982)) and their prior beliefs about similarity judgements (whether they believe thematic similarity is a genuine form of similarity) (Simmons and Estes, 2008, Exp. 3). Participants were more likely to exhibit thematic processing if they had a low need for cognition (implying that thematic processing requires less 'deep' processing) and if they believed that thematic processing was a genuine form of similarity. Mirman and Graziano (2012) found that preference for featural or thematic processing were also independent of age and education.

In conflict with the Dual-Process model of similarity judgements, it has been argued that thematic processing is an intrusion on, rather than a component of psychological similarity (Gentner and Brem, 1999; Honke and Kurtz, 2019). In this view, people can get a sense of thematic relatedness between two concepts and mistakenly allow this perception to influence their judgements in similarity tasks. Honke and Kurtz (2019) found that a preference for thematic matches in forced-choice tasks wanes over the course of an experiment (eventually leading to a preference for featural processing), is partly dependent on task instructions, and may be artificially inflated by design limitations such as the number of stimuli used and the size of participant samples. They interpret these results as characteristics of a thematic relatedness process that intrudes on similarity judgements, but does so only in certain circumstances. Whether thematic processing is a 'genuine' component of similarity or not, a person's perception of thematic relatedness *can* influence their similarity judgements.

3.3 The Dual-Process model of conceptual combination

According to Wisniewski's Dual-Process model (Introduced in Chapter 3), conceptual combination occurs via one of two cognitive processes that produces one of three kinds of combination (Table 3-1).

- Property-mapping and hybrid interpretations are the two possible outputs of a comparison and construction process. This process operates on the common and different features of a pair of concepts to identify suitable features to transfer between the concepts to create a new one.

- Relation-linking interpretations are the result of a scenario creation process that takes advantage of thematic relations between base concepts. Combinations are produced by placing the two base concepts into complementary roles bound by an external relation.

These two processes operate on the same dual-processes as those associated with the Dual-Process model of similarity judgements. Originally, the Structural Alignment model was used as an explanation for how property-transfer combinations occur in conceptual combination (Wisniewski, 1997a; Wisniewski and Markman, 1993). In reverse, knowledge of how people produce conceptual combination interpretations later influenced research into thematic similarity (Wisniewski and Bassok, 1999), giving rise to the Dual-Process model of similarity judgements (Chen *et al.*, 2014; Estes, 2003a, 2003b; Estes *et al.*, 2011; Wisniewski, 1997a).

Table 3-1 – Types of combinations and their processes in the Dual Process Model of conceptual combination (Wisniewski, 1997a).

General type	**Uses**	**Interpretation type**	**Description**	**Example**
Comparison & construction	Commonalities, alignable differences, nonalignable differences	Property interpretations	"a property of one combining concept is asserted of the other"	cactus frog is a spikey frog
		Hybrid interpretations	A combination of the constituents or a conjunction of the constituents	saw-scissors could be interpreted as a dual-purpose tool that both cuts and saws)
Scenario creation	Thematic relations	Relation-linking interpretations	"a relation is asserted between the concepts being combined"	robin-hawk could be interpreted as "a hawk that preys on robins" (Wisniewski and Middleton, 2002))

Evidence for the three types of interpretation (property-mapping, hybrid and thematic) comes from an inductive analysis of the interpretations produced by experiment participants when given a variety of stimuli pairs (Wisniewski, 1996). The three types of combination, and thus the two processes that give rise to them, differ in frequency of occurrence. In one experiment, relational combinations occurred 53% of the time, property-mapped combinations 41% of the time, and hybrids 1%, with 5% being classed as 'other' (Wisniewski, 1996). In another, relation-linking combinations occurred about 71% of the time and relational combinations 29% (Wisniewski and Love, 1998).

3.3.1 Comparison and construction

In the comparison and construction process, 'comparison' refers to the same mental structural alignment process described for similarity judgements in Section 3.2.1.2. The comparison process highlights differences that are suitable for transfer from one concept to another. These differences are passed to the construction process which creates a new property in the output concept based on the selected difference.

Evidence that the process involves 'comparison' rather than some other process comes from the finding that the frequency of hybrids increases as the perceived similarity of the base-concepts increases (Wisniewski, 1996). With increasing perceived similarity between two concepts, it becomes more likely that the pair of concepts will have enough common properties to facilitate hybrid interpretation. Thus, there needs to be some mechanism that can highlight the commonalities between concepts, i.e., 'comparison'.

The construction process accounts for conceptual change in property-mapping combinations. Conceptual change refers to the effect wherein properties of one concept are not transferred directly to another, but rather the modifier is thought to provide information for the 'construction' of a new property in the head concept. Wisniewski (1998) gives the example of interpreting a 'ZEBRA CLAM' as "a clam with stripes," (p.1330). They propose that the transfer of the property 'stripes' from the zebra to the clam does not involve a literal copy of the stripes, suggesting that the stripes of the new concept might differ in length, thickness or closeness. Rather, in the Dual-Process model, properties of modifier concepts are a source for newly constructed (or 'instantiated') properties in the head concept. The construction process is contrasted with a 'copy and paste' process in which abstract properties of one concept are simply moved as-is to the other (Wisniewski, 1998).

Evidence that the comparison process occurs via structural alignment (as opposed to any other algorithm) is a correspondence between alignable differences and property-mapping combinations. Property-mapping interpretations, but not relation interpretations, consist of alignable differences from the original pairs (Wisniewski and Markman, 1993). For example, given the pair CAR and MOTORCYCLE, the number of wheels is an alignable difference (since both have wheels but they differ in number), but a roof is a nonalignable difference since it is unique to the car. If structural alignment operates on the pair then the special nature of alignable differences (Section 3.2.1.2) should make alignable differences, but not nonalignable differences, available for transfer. This would make people more likely to interpret a 'car motorcycle' as 'a motorcycle with four wheels', rather than a 'motorcycle with a roof.

While the model proposes that alignable differences form the basis of property-based combinations, the selection of which specific difference(s) are transferred may be influenced by context (Wisniewski, 1997a), salience (Estes and Glucksberg, 2000), diagnosticity (Estes and Glucksberg, 2000; Wisniewski, 2000), and cue and category validity (Wisniewski, 1997a).

3.3.2 Relational combination: scenario creation

The scenario creation process produces relational combinations by placing two concepts into complementary roles in a scene in which the concepts are bound by a thematic relation. For example, the concepts 'car' and 'hammer' may be interpreted as a hammer for hitting cars (such as would be used by a panel beater to repair a car). Relational combinations occur when a plausible scenario can be found that binds the two concepts into separate but complementary roles in a relation. In this example, 'hammer' adopts the agent role, and 'car' adopts the recipient role in the 'hit' relation. In conceptual combination, the output of this process is a specialised type of hammer described in a scene with a car.

The scenario creation process is generally the same as the one described for similarity judgements. According to the Dual-Process models it operates on a schema-based slot and filler model of concept representation and takes advantage of thematic relations between concepts. Much like in the Dual-Process model of similarity, the algorithmic level of the scenario creation process is poorly specified.

3.3.3 The interplay of the two processes

The two processes that give rise to these combination types are proposed to be distinct (Estes, 2003b) and compete in parallel (Estes, 2003a). Wisniewski and Love (1998) show that the choice of which process wins out is associated with prior use of a process (Wisniewski and Love, 1998), whether a plausible thematic relation exists, and the 'alignability' of the base concepts (Wisniewski, 1997a). As introduced in Section 3.2.3, alignability describes the extent of potential alignment between two concepts. The influence of alignability and thematic relations means that, as per similarity judgements, different pairs of stimuli can be relatively more or less compatible with one kind of processing or the other.

Property-mapping combination tends to occur when facilitated by alignability or in the absence of a plausible thematic relation. Alignability facilitates property-mapping combinations in two ways.

- First, the comparison process needs a minimum degree of common relational structure to identify alignable differences for transfer. As more similar concepts tend to have a more extensive overlapping relational structure, similarity can thus be said to facilitate property-mapped combinations.
- Secondly, the transfer of property between concepts requires that the recipient concept has matching dependencies to accommodate the new property. For example, HAMMER FEATHER is unlikely to be interpreted as a feather for hitting things, since it lacks the mass and rigidity to facilitate such an action. Once again, more similar concepts are more likely to have matching dependencies and are thus more compatible with property-mapping combinations.

Wisniewski and Love (1998) (p.180) have also shown that property-mapping combinations predominate in the absence of a thematic relation.

For an individual to produce a relational combination they must determine that the two design concepts can plausibly play complementary roles in a thematic relation. The tendency to produce relational combinations is mostly independent of pair perceived similarity or alignability, with one exception. Concepts having many common properties may prevent relational combination since highly alignable concepts are more likely to occupy the same role in a relation (Wisniewski, 1997a, 2000).

No studies appear to have investigated individual differences in conceptual combination, but individual preferences for featural or thematic processing (Mirman and Graziano, 2012) may also influence the tendency for individuals to combine concepts through comparison & construction or scenario creation processes.

3.4 Discussion

A model of *design concept similarity judgements* would be valuable in the development of a model of *design concept combination* for two reasons. First, it would provide an independent test of Structural Alignment theory in a design context. If there is support for Structural Alignment in similarity but not in combination, then it can be inferred that Structural Alignment *does* occur as expected in design, but that alternative models for combination are required. If neither model is supported, then there would be two points of evidence indicating that there may be something unique about design cognition. Second, knowledge about designer cognitive processes would make an independently valuable contribution to knowledge. At present, there are no cognitive models of designer similarity judgements. An initial plan to use the Dual-Process model of conceptual combination as the basis of a model of design concept combination was going to involve the manipulation of similarity as an independent variable in an experiment. This is because similarity has a special role as a proxy for the determinant of processing type. Since it would be necessary to operationalise similarity anyway, it was deemed an efficient extension of the research to also conduct an empirical test of the model.

Similarity judgements are relevant to the study of conceptual combination because, as has been seen in Sections 3.2 and 3.3, accounts of both phenomena propose that they involve dual processes of comparison (via structural alignment) and scenario creation processes. This common processing basis has facilitated hypothesis generation about the relationship between the two processes (similarity judgements and conceptual combination). This originally enabled Wisniewski to provide a theoretical explanation for how property-transfer combinations occur in conceptual combination (Wisniewski, 1997a) and facilitated hypothesis generation (Markman and Wisniewski, 1997; Wisniewski and Markman, 1993). In reverse, knowledge of how people produce conceptual combination interpretations later influenced research into thematic similarity (Wisniewski and Bassok, 1999), giving rise to the dual-process view of similarity (Chen *et al.*, 2014; Estes, 2003a, 2003b; Estes *et al.*, 2011; Wisniewski, 1997a)

3.5 Summary

To model the cognitive processes involved in design concept combination, it was proposed that the Dual-Process model of conceptual combination could be used as the basis for a new model of design concept combination. In the process of developing a research design to test such a model, the scope of the research was extended to include the objective of modelling the cognitive processes involved in design concept similarity judgements. This is because knowledge about similarity judgements has, in the past, been used to gain knowledge about conceptual combination. Analogously, it was determined that knowledge about designer similarity judgements would aid in providing knowledge about design concept combination.

The main contribution of this chapter is an overview of the two existing cognitive models from the psychology literature that form the basis of the proposed models of designer cognition in Chapter 4. Section 3.1 presents background literature on human similarity judgements both generally and in design. This highlighted the lack of any cognitive models that represent designer similarity judgements. The middle of the chapter presented the Dual-Process model of similarity judgements (Section 3.2) and the Dual-Process model of conceptual combination (Section 3.3), both of which involve common algorithmic level processes. The discussion (Section 3.4) provides more detailed rationale for why similarity judgements are relevant to conceptual combination by drawing on the background literature and cognitive models.

4 PROPOSED MODELS AND RESEARCH DESIGN

In the previous chapter, two cognitive models were summarised from the psychology literature; the Dual-Process model of similarity judgements and the Dual-Process model of linguistic conceptual combination. Both models involve the same dual processes of (i) comparison via structural alignment, and (ii) scenario creation via slot filling. As outlined in the research methodology (Section 1.3.2.2), it was proposed that these same dual-processes may be involved in *design concept* similarity judgements and *design concept* combination. To this end, this chapter presents proposals for Dual-Process models of *design concept* similarity judgements (Section 4.1) and *design concept* combination (Section 4.1.3).

Just as the newly proposed models are based on the existing models outlined in Chapter 3, so too can the models be tested by the same methods that were used in the development of the existing models. The research design (Section 4.2) defines the scope of the investigation, which aspects of the proposed models are to be tested, the research questions and hypotheses used to test the models, considerations of validity and pre-requisite methodological developments. A research plan shows how the methodological requirements relate to the model-testing objectives (Objs 3 and 4) and a research map is provided that shows the chronological record of the work that was conducted (Section 4.2.6).

4.1 Target constructs and proposed models

4.1.1 Target constructs

To study cognitive processes, it is first necessary to define the kind of person to whom those cognitive processes belong and the situations in which those cognitive processes might occur. This enables an evaluation of the validity of the models as representations of designer processing. The relevant constructs are the characteristics of the designer, the stimuli they are processing, and the setting in which cognitive processes takes place.

The main population of interest are the people who contribute to the design and development of the products and systems that come to market, i.e., professional product design engineers. To develop methods and tools to aid these designers, it is necessary to gain knowledge about the cognition of this population. Thus, the ideal constructs of interest are professional engineering designers, making similarity judgements about, and combining, design concepts created by themselves or their colleagues, embedded in the early stages of the conceptual design process, taking place in the physical and social settings in which professional practice occurs such as design consultancies or engineering firms. However, the study of professional product design engineers in experimental settings requires a lot of resources.

An alternative population of interest is that of student or trainee designers. This population is relevant because it is during the education of the student that they learn to use many of the methods and tools that they will one day employ in industry. Moreover, the student design process has commonalities with professional design; student designers make similarity judgements about, and combine, design concepts, created by themselves and by their team members in the conceptual design phase of design projects. These commonalities provide a basis for some generalisation, and an advantage of studying students is that they are more readily available in university settings and require fewer resources to recruit for experimental research. However, the projects are typically shorter than those found in industrial settings, the design concepts are made by individuals with less experience and the physical setting may be a university building. Thus, students may be considered a proxy for professional engineering designers, with the caveat that knowledge about the cognition of student designers does not necessarily generalise to the cognition of professional designers.

For both professionals and students, similarity judgements and combination may be carried out on at least two inputs (design concepts), but there is no known upper limit to how many stimuli can be processed at once. Similarity judgements could be made on a group of design concepts (although this might be more aptly termed a judgement of variety) and it has been shown that designers can combine elements of numerous concepts to create new design concepts (Gonçalves and Cash, 2021). This is important, because the psychological models reviewed in Chapter 3 and the models of designer cognition proposed in this chapter are inherently constrained to the processing of pairs of stimuli.

4.1.2 The proposed Dual-Process model of design concept similarity judgements

The proposed cognitive model of design concept similarity judgements is shown in Figure 4-1. The model represents the cognitive processes through which a designer judges the similarity of a pair of design concepts. Specifically, it represents the situation in which a designer is presented with a pair of design concepts (input) and the instructions to provide a numerical similarity rating (goal) of how similar those concepts are (output). This perception of similarity is proposed to arise from one of two cognitive processes.

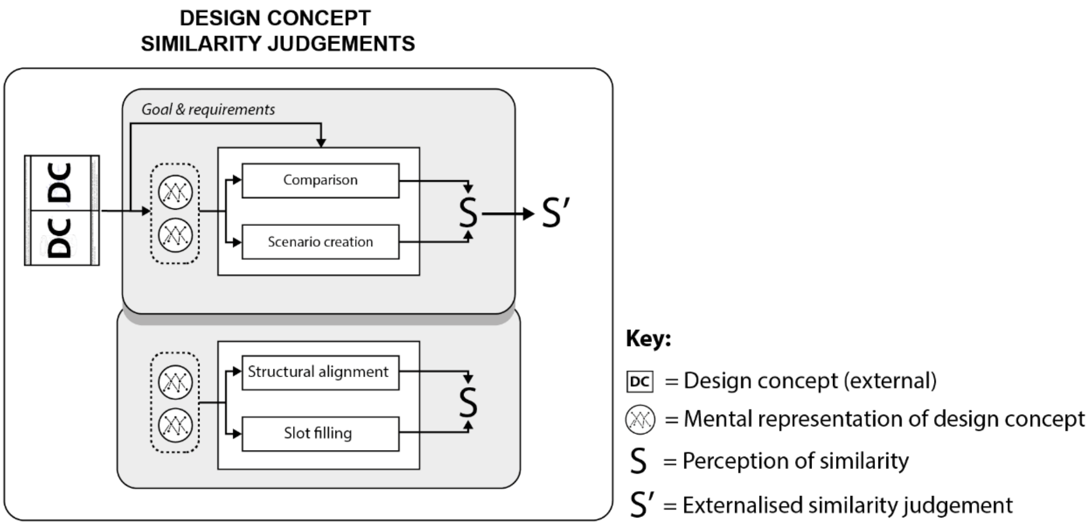


Figure 4-1 – The proposed cognitive model of design concept similarity judgements, showing the inputs (a pair of design concepts), output (rated similarity) and two cognitive processes at the computational and algorithmic levels.

Comparison via structural alignment. A comparison-based similarity process is proposed, in which the designer's perception of similarity arises from the comparison of the constituent features of two design concepts. Similarity is a function of the common and different features and the more common features that two design concepts share, the more similar they should be.

At the algorithmic level, it is proposed that this comparison process operates via the same processes as described in (Section 3.2.1.2). Specifically, it is assumed that design concepts are represented in the mind as structured representations and that comparison-based similarity judgements occur via a process of Structural Alignment. The comparison process involves the alignment of the relations of the two mental structured representations. Similarity is specifically a function of the commonalities, alignable differences and nonalignable differences between the two design concepts. It is further assumed that the structural alignment process adheres to the constraints of parallel connectivity and one-to-one mapping. The model makes four predictions about the relationship between similarity and commonalities and differences.

1. A designers perception of similarity is a function of commonalities, alignable differences and nonalignable differences, and alignable differences should be a greater influence on similarity than nonalignable differences.
2. Pairs of design concepts with increasing perceived similarity should also have more commonalities
3. Pairs of design concepts with many commonalities should also have many alignable differences.
4. Alignable differences should be easier to identify than nonalignable differences.

Scenario creation via slot filling. A scenario-creation process is proposed, in which the designer's perception of the similarity is influenced by the perception of a thematic relation between the two design concepts. Through this process, two design concepts may be perceived as being more similar if the designer thinks that they 'go together' or can envision their placement into complementary roles in a system or scenario in which they are bound by an extrinsic, complementary relationship.

No explicit claims are made about the algorithmic level of the scenario creation process as this is beyond the scope of this thesis.

4.1.3 The proposed Dual-Process model of design concept combination

The proposed cognitive model of design concept combination is shown in Figure 4-2. The model represents the cognitive processes through which a designer combines a pair of design concepts to create a new design concept. Specifically, it represents a scenario in which a designer is presented with a pair of design concepts (input) and the instructions to combine those design concepts (goal) to create a new one (output). According to this model, combination occurs via one of two processes that each lead to different types of design concept.

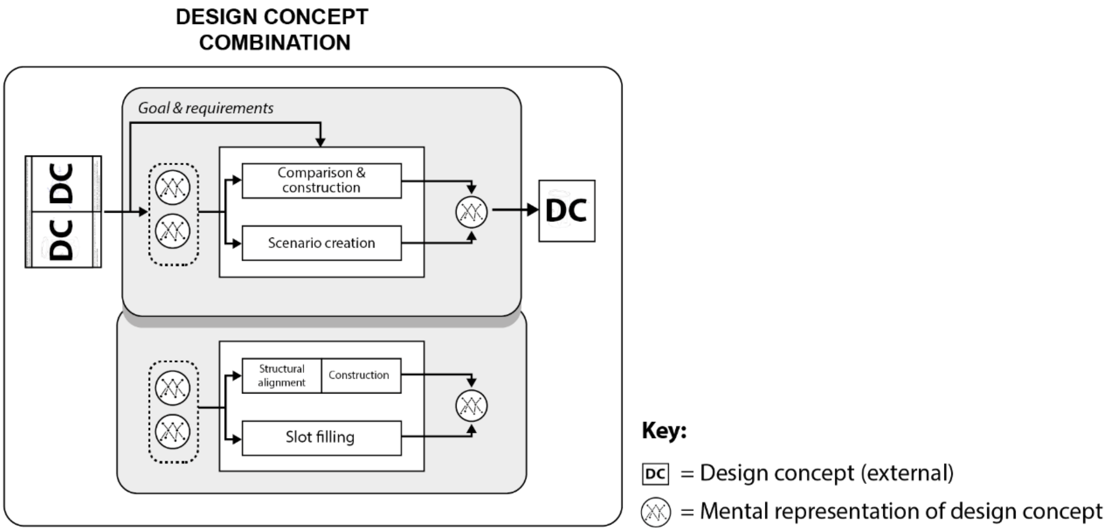


Figure 4-2 – The proposed cognitive model of design concept combination, showing the inputs (a pair of design concepts), output (rated similarity) and two cognitive processes at the computational and algorithmic levels.

Comparison (via structural alignment) and construction. The 'comparison and construction' process comprises two processing steps[11]. The comparison process is the same as the one described in the Dual-Process model of design concept similarity judgements. A structural alignment algorithm places the relational structure of the two design concepts into alignment, making alignable differences (but not nonalignable differences) more salient, and thus more available for transfer between the two design concepts. These salient differences are then used by the construction process to produce new features in the combined concept.

[11] The comparison and construction processes are illustrated as two serial processes (Figure 4-2). This is done to draw a conceptual distinction between them but these processes may well be indissociable. It may be possible to further decompose these processes into mechanisms responsible for selecting what specific differences are transferred between design concepts (Wisniewski and Middleton, 2002), but this is beyond the scope of the current model.

A comparison and construction process would result in a single design artefact that contains a feature (or features) from both base concepts. The construction process is included to account for the expected presence of conceptual change. That is, rather than features being 'transferred' between concepts like rigid physical components moving spatially between products, features in the base concept act as the basis for *new* features in the combined concept with modified properties.

Scenario creation. The proposed scenario creation facilitates combination by placing two design artefacts into a scenario or system in which they play complementary roles in an extrinsic relation. As per the Dual-Process model of design concept similarity judgements, this process is assumed to operate via a slot-filling algorithm but this is beyond the scope of the research in this thesis. The distinguishing feature of the design concepts created through the scenario creation process is that design artefacts from each base concept are both identifiable in the combined design concept. Minor intrinsic modifications may take place to accommodate the change of role(s), but otherwise, the base concepts should be recognisable in the resulting scenario.

Four claims are made that relate to the relationship between similarity and combination type. They are derived from the Dual-Process model of conceptual combination as described in Section 3.3.3. They are:

1. Similarity facilitates featural combinations.
2. Relational combinations occur in the presence of plausible thematic relations.
3. Featural combinations predominate in the absence of a thematic relation.
4. Relational combinations are inhibited by very high base-concept similarity.

These claims relate to computational level processing and do not allow for falsifiable tests since the claims of Wisniewski's model are verbal and lack quantitative criteria for falsification.

4.2 Research design

The research design includes an overview of the planned studies, considerations about validity, a list of pre-requisite methodological developments, objectives to guide the empirical research and a research map that shows the chronology of the work that was conducted.

4.2.1 Scope of the research and the use of research questions and hypotheses

The scope of the research is limited to specific parts of the proposed cognitive models. As shown in Figure 4-3, the intention was to investigate the computational level of each model (1, 3) to determine whether there was evidence for dual-processes of Structural Alignment and scenario creation. Then, if there was evidence for a Structural Alignment process (2, 4) additional hypotheses would be subjected to falsification testing. This forms a stage-gate, where progress to the algorithmic level is dependent on the data pertaining to the computational level. The algorithmic-level slot-filling algorithms thought to be involved in each scenario creation process were not addressed, both because they are not concretely defined in the psychological literature, and because their inclusion would be beyond the possible scope of the research project.

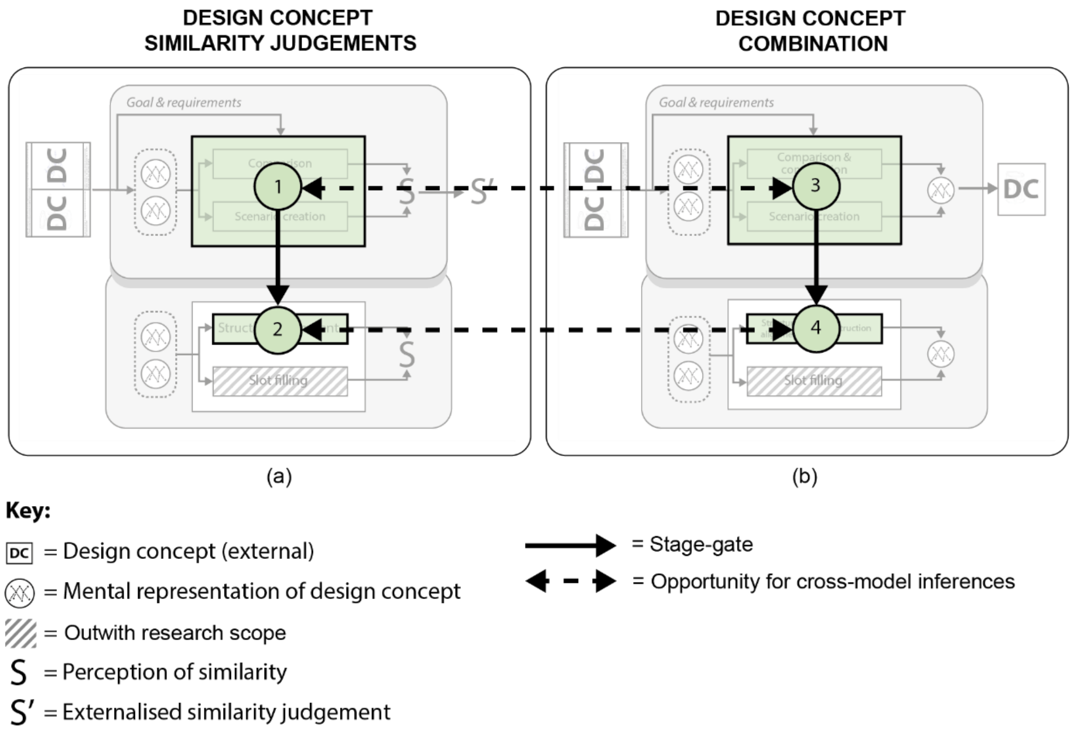


Figure 4-3 – The scope of the empirical research showing what cognitive processes are to be investigated and at what level of description (computational or algorithmic).

4.2.2 Research questions and hypotheses

4.2.2.1 Design concept similarity judgements

Three research questions were established to look for evidence that the similarity ratings were produced via either of the dual processes (designated RQ – S, below). Following Wisniewski and Bassok (1999), if designers make similarity judgements of design concepts via comparison and scenario creation processes then it would be expected that their self-report justifications for their similarity ratings would contain evidence of features or thematic relations. Similarity judgements carried out via comparison should refer to the commonalities or differences between the concepts and similarity judgements carried out via scenario creation should refer to the two design concepts and the extrinsic relation between them. The first two research questions were set to establish whether this was the case and, if so, how prevalent each type of explanation was.

RQ – S1: *Do designer explanations for their similarity ratings contain (i) commonalities or differences, or (ii) thematic relations?*

RQ – S2: *What is the relative prevalence of each explanation type?*

If the participant responses are indicative of comparison and scenario creation then, analogously to the Dual-Process model of similarity judgements (Section 3.2.1.2), the prevalence of each explanation type may be influenced by the relative similarity of the base concepts or the presence of a stimulus compatibility effect. This led to the third research question.

RQ – S3: *What is the relationship between similarity and explanation type?*

Five hypotheses were adopted from prior research (Markman and Gentner, 1993b, 1996) (designated H – Sx, below). If designers make similarity judgements exclusively or predominantly via structural alignment processing, it would be expected that the predictions of the Structural Alignment model of similarity judgements would also apply to design concept similarity judgements. To test this,

H – S1a: *Similarity should increase as a function of commonalities and decrease as a function of differences and commonalities should influence similarity more than differences.* Regression analyses should show that similarity can be predicted by the number of listed commonalities (positive regression coefficient) and total differences (negative regression coefficient). The unstandardized regression coefficient for commonalities should also be larger than that for differences.

H – S1b: *Alignable differences should be more important in evaluating similarity comparisons than nonalignable differences (Markman and Gentner, 1996).* An extension of H-S1a. This may be demonstrable using regression analysis. Similarity should be predictable by the number of listed commonalities (positive regression coefficient) and alignable and nonalignable differences (negative regression coefficients). The unstandardized regression coefficient for alignable differences should be greater than that for nonalignable differences.

H – S2: *Similar concepts should be associated with an increased number of commonalities and dissimilar concepts should be associated with a decreased number of commonalities.* This should be evident in positive correlations between rated similarity and the number of listed commonalities.

The next two hypotheses are unique to the Structural Alignment model and are only relevant if one assumes that concepts are represented in the mind as structured representations and there is a distinction between alignable and nonalignable differences.

H – S3: *There should be a numerical link between commonalities and alignable differences.* This should manifest as a positive correlation between the number of listed commonalities and the number of listed alignable differences. A secondary prediction based on this is that because there should be a positive correlation between similarity and commonalities (Hypothesis 0b), there should also be a positive correlation between similarity and alignable differences.

H – S4: *Alignable differences should be more numerous than nonalignable differences.* This should manifest as a statistically significant increase in the number of listed alignable differences versus the number of listed nonalignable differences.

Strictly, the Structural Alignment model makes predictions H-S1b, but H-S1a is also included because it is a prediction common to feature-set and feature-alignment models. That means that both the Contrast and Structural Alignment models should make this prediction. It is included as a first-pass check of the applicability of a featural model of design concept similarity judgements.

4.2.2.2 Design concept combination

Four research questions were established to provide an initial test of the applicability of the Dual-Process model of design concept combination (designated RQ – C, below). Following research in linguistic conceptual combination (Wisniewski and Bassok, 1999) and design (Nagai *et al.*, 2009), if designers combine design concepts via processes of (i) comparison and construction and (ii) scenario creation, then it would be expected that the combined design concepts would contain evidence as to how they were produced. Combination carried out by comparison and construction should result in a single design artefact that contains features from both base concepts. Combination carried out by scenario creation should result in a scenario or system of artefacts related by some extrinsic relation. This leads to the first two research questions.

RQ – C1: *What types of combinations do designers produce?*

RQ – C2: *What is the prevalence of each combination type?*

These research questions were addressed through an abductive coding process. The three interpretation types from the Dual-Process model (Table 3-1) were used as the basis for coding types of combined design concepts, but the codes were developed through iteration to accommodate the characteristics of the design concepts.

The third research question concerns the relationship between similarity and combination type. A measure of the relationship between similarity and combination type can be used to assess the four claims of the Dual-Process model. Similarity is a key outcome measure in research associated with Wisniewski's Dual-Process model, because it is theoretically related to the underlying properties of alignability and thematic relatedness, which are more difficult to measure directly than similarity. This leads to the third research question:

RQ – C3: *What is the relationship between concept pair similarity and the type of combined concept?*

A final question concerned the potential role of combination difficulty as a confounding variable. The concepts of ease and difficulty can be found in the conceptual combination literature, although it is not clear to what degree this language is intended to refer to a direct influence of difficulty. This is highlighted in two examples (bold emphasis added). When describing the tendency for individuals to produce hybrid combinations, Wisniewski notes that "it is very **difficult** to interpret a *drill pamphlet* as a hybrid, since it would require both the function of a drill and a pamphlet and these properties conflict." (Wisniewski, 1997, p.175). When describing the competition between the dual processes, it is said that "Since it is **easier** to align representations with similar structure and to find their commonalities and differences, the comparison/construction process "wins" over scenario creation even though there is a plausible interpretation involving a scenario." (Wisniewski, 1997, p.178). Since it is not clear what the role of difficulty is in combination, the fourth research question is intended to pre-empt any potential confounding effect of similarity on combination type.

RQ – C4: *What is the relationship between concept pair combination difficulty and the type of combined concept?*

Two hypotheses were derived from the Dual-Process model of conceptual combination (designated H – C, below). They are based on the empirical evidence that Wisniewski (1997) uses in support of the Dual-Process model of conceptual combination. The first hypothesis concerns the relationship between hybrid combinations and stimulus similarity.

H – C1: *Relatively more similar pairs of design concepts should be associated with a greater proportion of hybrid combinations.* In conceptual combination, hybrid combinations contain many features from both base concepts (Table 3-1). The frequency of hybrid combinations increases with similarity because it becomes increasingly likely that there will be enough common features to facilitate hybrid combinations. This hypothesis can be tested via correlational analyses of base-concept similarity and the proportion of hybrid combinations. Crucially, this implies that it is possible to distinguish a 'hybrid' combination from any other kind and is thus dependent on the outcome of RQ-C1.

The second hypothesis concerns the relationship between alignable differences and feature-mapping combinations. They are combinations in which a new feature from one base concept has seemingly been transferred to the other. The prediction is that such combinations exist in design concept combination, and if a structural alignment process is involved, the special nature of the alignable differences should mean that alignable differences, not nonalignable differences, should be transferred between concepts. Once again, this hypothesis is dependent on the outcome of RQ-1 and the ability to distinguish between 'feature mapping' and 'hybrid' combinations.

H – C2: *There should be a correspondence between alignable differences and featural combinations, but not alignable differences and relational combinations.*

Following Wisniewski and Markman (1993), it would be expected that there would be a statistical covariation between consensus alignable differences (those listed by more than a threshold number of participants) and the equivalent of 'property' combinations but not 'relational' combinations.

4.2.3 Planned studies

To address the research questions and hypotheses, a series of correlational quasi-experiments were planned. Figure 4-4 shows the elements of the planned studies (white rounded rectangles) and how they relate to the elements of each proposed model (represented by the green rounded rectangles and their grey backdrops at the bottom of the figure, corresponding to Figure 4-3). The vertical arrows show how each element of the planned experimental research contribute to addressing research questions and hypotheses.

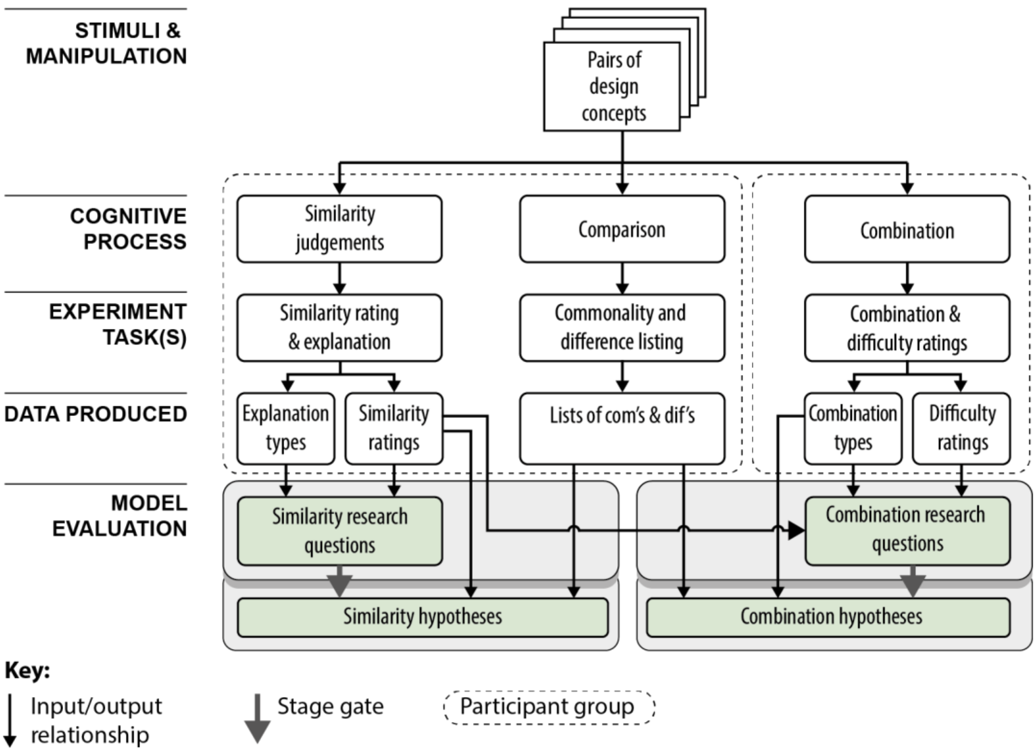


Figure 4-4 – Elements of the research design and their relationship to the proposed models.

The study elements in Figure 4-4 are as follows.

- **Stimuli and manipulation -** All studies involved the presentation of pairs of design concepts that vary in relative similarity. Using the same stimuli in parallel similarity and combination experiments facilitates the drawing of inferences across models.
- **Cognitive processes and tasks -** Three cognitive processes are elicited: similarity judgements, comparison and combination. The computational level of each model can be evaluated by investigation of that process alone (i.e., similarity or combination). Evaluation of the algorithmic levels also requires the elicitation of

comparison processing. The predictions made by Structural Alignment Theory in each of the proposed models concern the relationships between the outputs of a task that is thought to involve comparison (i.e., similarity judgements or combination) with the outputs of a task that inherently elicits comparison (commonality and difference listing). An additional difficulty-rating task is used to provide a measure of combination difficulty. Supplementary interviews and observations were used to aid in methods development and are outlined in Chapter 5.

- **Participant group allocation -** The cognitive processes are elicited in a mixed within-groups and between-groups design. One group of participants makes similarity judgements and listed commonalities and differences, the other group of participants carries out combination.
- **Tasks and data -** Similarity processing is engaged by a similarity rating task and an additional self-report task that asks participants to explain their rationale for the similarity ratings. Comparison is engaged by tasking participants with listing the commonalities and differences of pairs of concepts. Design concept combination is engaged by a combination task that asks the participant to combine concepts to create new ones. Some of the data produced needs to be coded before it can be used.

The study elements shown in Figure 4-4 are all based to some extent on prior studies in the psychology literature but require varying amounts of adaptation to be compatible in a design context. Requirements were specified for stimuli, coding schemes and assumptions about design concept knowledge representation.

4.2.4 Pre-requisite methodological contributions

Three methodological developments were necessary to facilitate the empirical research. These were (i) stimuli, (ii) coding schemes and (iii) appropriate theoretical assumptions about how design concepts are represented in the mind (the ‘representing world’).

Pairs of design concepts as experiment stimuli. There was a need to create pairs of design concepts that vary in relative similarity. This was necessary because the hypotheses associated with the proposed models both make predictions about covariation between similarity and some other dependent variable.

A minimum requirement set for pairs of design concepts that varied in relative similarity as validated by human similarity ratings. No constraints were placed on the nature of this variation and so it may have been achieved by e.g., high- and low- similarity groups or pairs spanning a continuum of similarity. Nor were constraints placed on the method to be used for creating the design concepts.

Coding schemes for the commonality and difference listing task and the design concept combination task. There was also a need to develop two coding schemes. One was for coding alignable and nonalignable differences in the commonality and difference listing task. The other was for coding types of combined concepts in the design concept combination task. This second coding scheme was influenced by assumptions about how design concepts are represented in the mind. Three requirements were set for such a scheme:

- Must describe the causal relational structure of design concepts in a format compatible with structured representations.
- Needs to represent designer knowledge.
- Must demonstrate utility in describing design concepts before and after they have been combined and must do so in a way that facilitates the interpretation of (i) feature-mapped and (ii) relation-linked combinations.

4.2.5 Validity

Validity is the property of an inference that describes the extent to which it can be taken as true based on the given evidence (Shadish *et al.*, 2002). In the present research, inferences are knowledge claims derived from the answers to the research questions and tests of the hypotheses. To maximise the validity of the inferences, steps were taken to mitigate threats to validity (reasons as to why inferences can be wrong) both when operationalising the constructs under investigation and when establishing procedures for analysing the data (Chapters 7 and 8). This was done by considering four types of validity (Table 4-1), adopted from Shadish *et al.* (2002).

Table 4-1 - Four types of validity (Shadish *et al.*, 2002)

Type of validity	The validity of inferences about…
Statistical conclusion validity	… the covariation between two variables
Internal validity	… whether observed covariation reflects a causal relationship
Construct validity	… the higher-order constructs that operationalisations are intended to represent
External validity	…whether the causal relationship holds over variations in study characteristics

The first two of types of validity, statistical conclusion validity and internal validity, concern the claims of whether an experimental manipulation caused the measured outcome. The other two, construct validity and external validity, concern claims about generalisability. Generalisability is important in (quasi-) experimental research owing to the inherent disconnect between the constructs under investigation (similarity and combination cognitive processes) and the particulars of the experiments used to study them. To bridge the gap between experiment and application (i.e., education or practice) it is necessary to generalise the findings beyond the particular study characteristics employed in these studies.

Statistical conclusion validity refers to inferences about the covariation between the independent and dependent variables. Covariation is used in the general sense of a relationship between two variables that may be identified through e.g., correlation, regression or group difference analyses. The two components of statistical conclusion validity are (i) whether there is covariation, and (ii) how strong that covariation is. Threats to the first component stem from factors that may contribute to Type I (false positive) and Type II (false negative) errors. Threats to the second relate to factors that may cause an over or understatement of the magnitude of covariation.

Internal validity refers to whether the change in a dependent variable (DV) can be attributed to a causal effect from the independent variable (IV). There are three requirements for a cause-effect relationship.

- Covariation – a demonstration of a relationship between the IV and DV.
- Precedence of a cause – a change in the dependent variable should be observed after the manipulation of the independent variable.
- Nonspuriousness – Alternative explanations for the observed covariation must be refutable.

Threats to internal validity are reasons why an inference that a relationship is causal may be incorrect, i.e., any violation of the three requirements above. Valid covariation is addressed by statistical conclusion validity, meaning that internal validity is dependent on statistical conclusion validity. The precedence of cause requirement is not met because of the correlational aspects of the research design. For example, the hypotheses associated with the proposed model of design concept similarity judgements make predictions about associations between similarity ratings and the number of commonalities that designers can list for a pair of design concepts. Although the independent variable (relative *a priori* similarity) was manipulated before the designers completed their tasks, the variables used in the analyses (similarity ratings and number of commonalities) were produced at the same time, are conceptually related, and cannot be disentangled.

The research design accounts for several potential threats to internal validity. Threats via instrumentation (changes in the delivery of the experiment) were addressed in conjunction with experimenter threats by standardising experimental procedures and task instructions. Threats from boredom or fatigue were addressed by randomising the stimuli presentation order. Testing effects (prior exposure to an experiment can affect performance in subsequent experiments) were prevented by making sure no participants took part in more than one experiment. Similar to testing effects, there was a risk that tasks eliciting processes with common cognitive processes could introduce priming effects. The steps taken to mitigate these effects are noted in Section 5.1.4.

Construct validity refers to the ability to make inferences from specific operationalisations to the more general constructs that they represent. Five study characteristics relevant in this thesis are the participants, the setting in which the study takes place, the stimuli, the variables that are manipulated, and the outcome measure. Good construct validity benefits from good construct explication, i.e., clearly defined constructs of interest and the selection of appropriate operationalisations for those constructs. As such, threats to construct validity may be any factor that impedes the match between construct and operation. Threats to construct include poor construct explication, confounds through overrepresentation of the construct (the operation captures more variables than intended), or bias from underrepresentation of the construct.

External validity refers to the generalisability of a (causal) relationship across variations in study characteristics. External validity differs from construct validity in that it refers to the *relationship* between the two operationalised variables (i.e., between manipulation and measured outcome) independent of whether those operationalisations reflect the intended constructs. A subtype of external validity is **ecological validity,** the extent to which causal relationships hold in real-life settings (Andrade, 2018). Inferences can have good external validity but poor ecological validity. For example, if inferences about the cognition of student designers in experimental settings generalised well across participants with different demographic backgrounds, participating in studies with different stimuli and different measures, the inference may be said to have high external validity. However, if these results do not generalise to professionals, or to cognitive processing taking place in naturalistic design environments, the inferences may have poor ecological validity. Shadish *et al.* (2002) frame threats to external validity in terms of interactions between the causal relationship and study characteristics. An interaction is deemed a threat if it is sufficiently large that it might change the direction of a relationship[12].

4.2.6 Research plan

The research plan (Figure 4-5) shows how the two methodological requirements (Section 4.2.4 relate to the two model-testing objectives (Obj. 3 and 4). Stimuli are required before either of the proposed models can be evaluated. Coding schemes will be required for analysing the Structural Alignment hypotheses associated with the similarity model, and the research questions associated with the combination model. The research questions and hypotheses are summarised in Table 4-2.

[12] Shadish *et al.*, (2002) note that an alternative conceptualisation of a threat is one that threatens the consistency of an effect size. This is not deemed relevant in this thesis since the hypotheses make predictions about the direction of relationships but not their size.

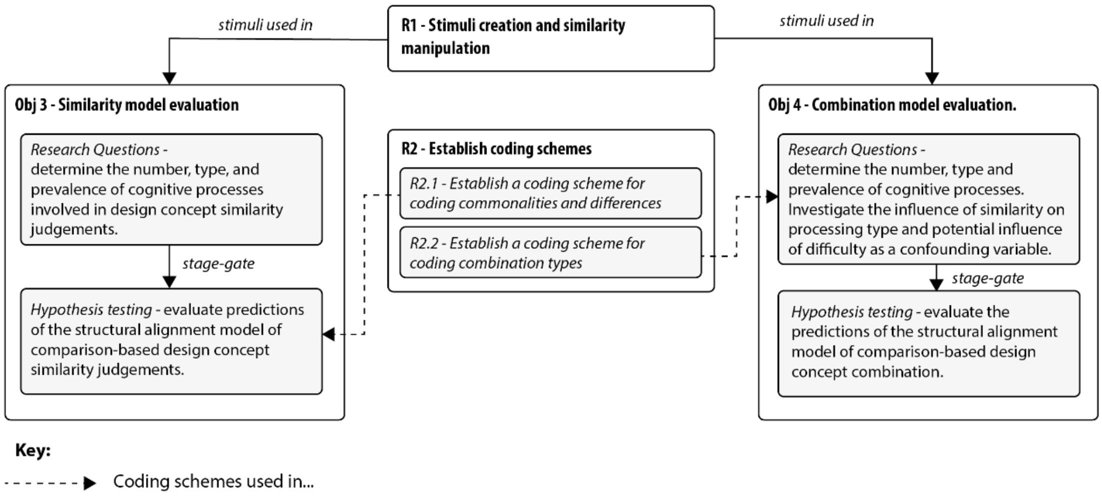


Figure 4-5 - The research plan

- **R1 - Stimuli creation and independent variable manipulation.** Create pairs of design concepts of varying similarity.
- **R2 - Establish coding schemes**. Coding schemes are needed to (i) classify types of differences from the commonality and difference listing task and (ii) classify types of combinations from the design concept combination task.
  - Establish a coding scheme for coding commonalities and differences
  - Establish a coding scheme for coding combination types
- **Obj. 3 - Similarity model evaluation.** Evaluate the proposed Dual-Process model of design concept similarity judgements.
  - *Research questions* – determine the number, type, and prevalence of cognitive processes involved in design concept similarity judgements.
  - *Hypothesis testing* – evaluate predictions of the Structural Alignment model of comparison-based design concept similarity judgements.
- **Obj. 4 - Combination model evaluation** – Evaluate the proposed Dual-Process model of design concept combination.
  - *Research questions* – determine the number, type and prevalence of cognitive processes. Investigate the influence of similarity on processing type and the potential influence of difficulty as a confounding variable.
  - *Hypothesis testing* – evaluate the predictions of the Structural Alignment model of comparison-based design concept combination.

Table 4-2 – Summary list of all research questions and hypotheses for empirical research

Ref.	Research question or hypothesis
Obj. 3 – Similarity model evaluation	
Research questions, computational level	
RQ – S1	Are explanations for design concept similarity ratings indicative of (i) a comparison process and (ii) a scenario creation process?
RQ – S2	What is the relative prevalence of each explanation type?
RQ – S3	What is the relationship between concept pair similarity and the type of similarity-explanation?
Hypotheses, algorithmic level	
H – S1a	Similarity should increase as a function of commonalities and decrease as a function of differences and commonalities should influence similarity more than differences.
H – S1b	Alignable differences should be more important in evaluating similarity comparisons than nonalignable differences (Markman and Gentner, 1996).
H – S2	Similar concepts should be associated with an increased number of commonalities and dissimilar concepts should be associated with a decreased number of commonalities.
H – S3	There should be a numerical link between commonalities and alignable differences.
H – S4	Alignable differences should be more numerous than nonalignable differences.
Obj. 4 – Combination model evaluation	
Research questions, computational level	
RQ – C1	When designers combine design concepts to create new ones, are the resulting design concepts indicative of (i) a comparison process and construction process and (ii) a scenario creation process?
RQ – C2	What is the relative prevalence of each explanation type?
RQ – C3	What is the relationship between concept pair similarity and the type of combined concept?
RQ – C4	What is the relationship between concept pair combination difficulty and the type of combined concept?
Hypotheses, algorithmic level	
H – C1	Relatively more similar pairs of design concepts should be associated with a greater proportion of hybrid combinations.
H – C2	There should be a correspondence between alignable differences and feature-mapped combinations, but not alignable differences and thematic combinations.

4.2.7 Research map: chronological phases and chapter structure

Figure 4-6 shows a map of how the research was actually carried out in practice, highlighting the chronology of the work and showing where the work is reported in the thesis. The map can be read from two perspectives: (i) according to the chapters in which the research is reported, represented by the grey background boxes, and (ii) in terms of the three chronological phases of research read top to bottom. The map is explained in terms of its content, the three phases of research and the correspondence between chapter contents and the research plan.

Content. The map represents the following elements:

- **Phases.** The research was carried out in three sequential phases, denoted by the black vertical 'phase indicator' bars at the left of Figure 4-6.
- **Assumptions about mental concept representation**. Each phase of empirical research was conducted based on different assumptions about the design artefact knowledge representation.
- **Methods**. Grey boxes denote the experimental methods or stimuli creation methods. The experiments are described using the nomenclature Sim-Px or Combo-Px. 'Sim' and 'Combo' denote experiments associated with the models of similarity and combination respectively and 'Px' refers to each phase of research, where x is 1, 2 or 3.
- **Stimuli**. The pairs of design concepts that are presented to participants in the similarity and combination experiments. There are three stimuli sets in total, a different one for each phase.
- **Evaluations**. In phases 1 and 2, the stimuli creation and experimental methods were evaluated to determine whether they were fit for purpose.
- **Analyses and results**. The data from the similarity and combination experiments were analysed to answer research questions and test hypotheses associated with the similarity and combination models.

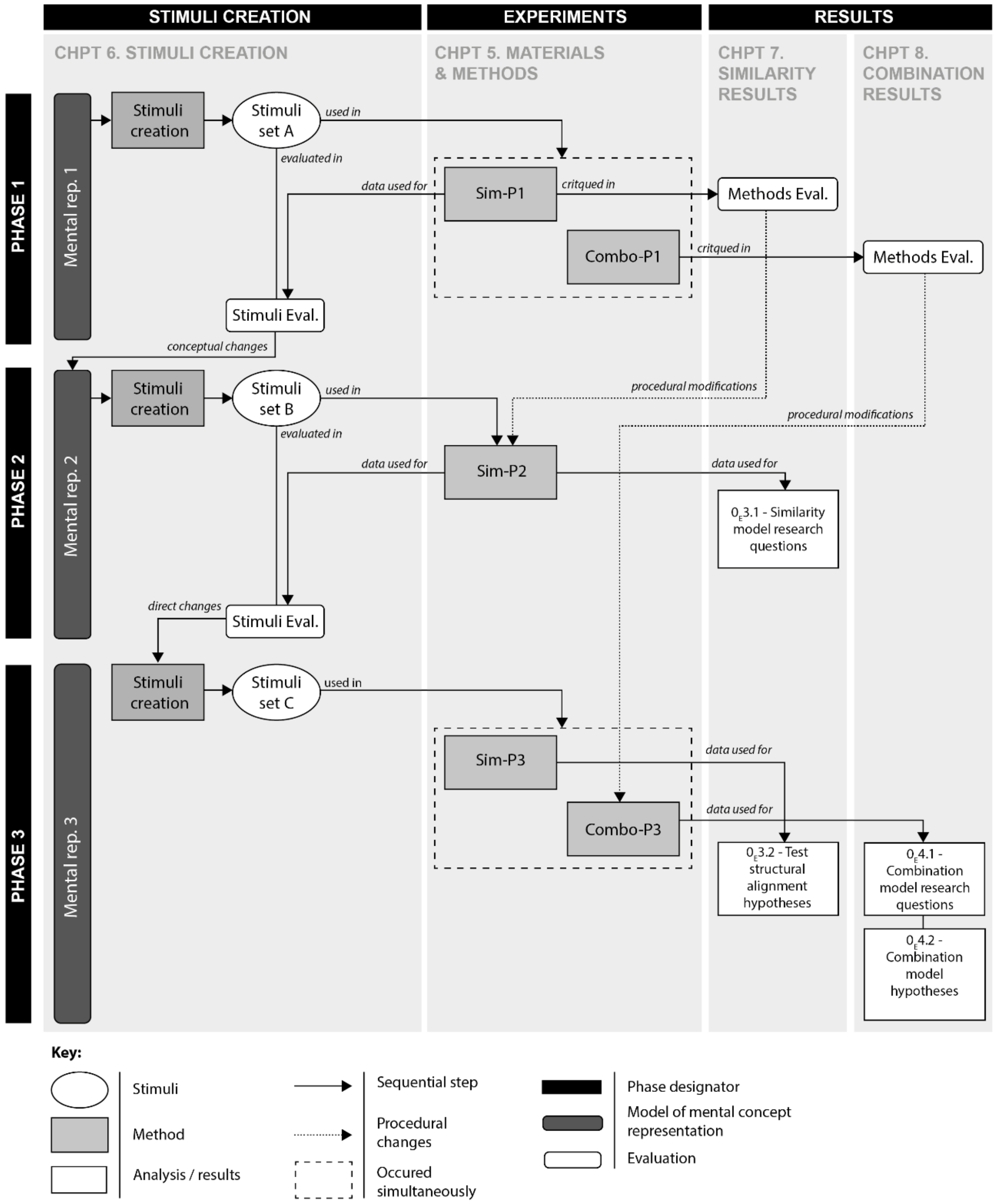


Figure 4-6 – Research map showing how the chronology of the empirical research and how the work is reported across five chapters.

Phases. Each phase of empirical research began with a set of assumptions about how design concepts are represented in the mind. This influenced the stimuli creation and interpretation of the results.

- **Phase 1** – The first phase involves pilot studies for stimuli creation and the similarity and combination experiments. It was initially assumed that design concepts were represented in the mind by lists of function, behaviour, and structure variables. Stimuli set A is created based on this assumption through manipulating the number of matching and mismatching F, B and S variables in an attempt to provide *a priori* control over the similarity of the stimuli. These stimuli were then used in Sim-P1 and Combo-P1. The key outputs from this phase were refinements to the methods.

- **Phase 2** – The second phase focused solely on the similarity experiments. It was assumed that design concepts can be represented by lists of 'Purpose', 'Sub-purpose', 'Function' and 'Means' variables and that they can be categorised into hierarchical design concept ontologies. Stimuli set B was created based on this assumption through a process of design concept categorisation and coding and the stimuli were used in Sim-P2. The similarity manipulation and experiment procedures are evaluated (Eval 2). The analysis of the data from Sim-P2 are presented in Section 7.1 to address the research questions associated with the similarity model (Obj. 3). Minor procedural changes are proposed for Sim-P3.
- **Phase 3** – The final phase addressed the hypotheses from the similarity model and addressed all aspects of the combination model. Stimuli set C was created based on minor modifications to stimuli set B. The stimuli are used in Sim-P3 and Combo-P3. The results from Sim-P3 are reported in Section 7.2 and the results from Combo-P3 are reported in Chapter 8 to complete the research for Objs. 3 and 4, respectively.

Chapters and objectives. The empirical research is reported as follows:

- **Chpt. 5 – Materials and methods**. This chapter describes the similarity and combination experiment methods across three phases.
- **Chpt. 6 - Stimuli creation**. Presents the research done to create the stimuli used in each phase of research (R1).
- **Chpt. 7 – Design concept similarity judgements: results**. Presents the pilot study (Sim-P1) and the results of Sim-P2 and Sim-P3.
- **Chpt. 8 – Design concept combination: results.** Presents the results of Combo-P3, thereby addressing the research questions and revealing that the planned analysis of the combination hypotheses (H-C1 and 2) is not possible owing to unforeseen differences between conceptual combination and design concept combination.

4.3 Chapter Summary

Based on two existing models of similarity judgements and conceptual combination (Chapter 3), new models of design concept similarity judgements and design concept combination have been proposed. Both models involve common cognitive processes of (i) comparison via structural alignment and (ii) scenario creation via slot-filling. A research design was developed to test the models by using quasi-experiments to answer research questions and test hypotheses. The scope of the empirical research is limited to the testing of the computational level of each model (number, type and prevalence of each process) and the algorithmic level of the structural alignment algorithms. A mixed within-groups and between-groups research design is presented, comprising correlational quasi-experiments. Considerations concerning the validity of the outcomes of the empirical research are discussed (Section 4.2.5). A plan for the empirical research is provided (Section 4.2.6), accompanied by a research map (Figure 4-6) that shows the empirical research that was carried out, showing both the chronological order of the work and the chapters in which it is presented.

5 MATERIALS AND METHODS

This chapter presents the materials and methods for all of the studies conducted in the thesis. As shown in the research map (Figure 4-6) there were two parallel streams of research (similarity and combination) across three phases of research, and experiments conducted in the same phase used the same stimuli. Section 5.1 presents general materials and methods that are relevant to multiple experiments; this helps avoid repetition in subsequent sections. The specific procedures for the three similarity experiments are presented in Section 5.2 and for the two combination experiments in Section 5.3. Details about the development and reliability of coding schemes are presented in these two sections. The experiments reported in this chapter are referenced in Chapters 6 (stimuli creation), 7 (similarity experiments results) and 8 (combination experiments results).

5.1 General materials and methods overview

The materials and methods presented in this section are those relevant to multiple experiments. They are presented here and cross-referenced in Sections 5.2 and 5.3 to reduce repetition.

5.1.1 Participant population and sampling

As noted in Section 4.1.1, the main population of interest is that of professional designers, but student designers can act as a proxy for professionals with the caveat that the extent of the generalisability between the two populations cannot be known without further experimentation. Populations of students were used in the research owing to pragmatic constraints. That is, they were more readily available and cheaper than recruiting professionals. Table 5-1 shows which populations were used for each experiment. The two populations were:

- **Population 1 – Engineering/design familiar**
  University students who had taken at least a one-semester general design class in the context of mechanical engineering or product design engineering. For example, undergraduate or postgraduate students with a mechanical engineering or product design degree.
- **Population 2 – Product Design Engineering competent**
  Undergraduate (final semester of 2nd year or any year above) or postgraduate students of Product Design Engineering (PDE), up to individuals who had graduated from a PDE degree within the 2 years before participation.

Population 2 is a subset of population 1 where the participants are more homogeneous in educational background and have more extensive design experience. Population 1 reflects the most easy to access participant pool available to the researcher that were deemed to have sufficient design experience to complete the similarity experiments. Population 2 reflects a more homogeneous population with a greater extent of design experience. The more extensive experience was deemed necessary for completing the combination experiments and the increased homogeneity was beneficial for minimising experimental confounds stemming from domain expertise, but population 2 was smaller and more difficult to recruit from than population 1.

Table 5-1 - Participant populations and the experiments they were used in.

	Phase 1		**Phase 2**	**Phase 3**	
	Sim1	**Combo1**	**Sim 2**	**Sim 3**	**Combo 3**
Engineering/design familiar	✓		✓		
PDE competent		✓		✓	✓

Participants were drawn from either the Product Design Engineering courses at Strathclyde University or Glasgow University/Glasgow School of Art and were recruited through multiple recruitment channels:

- Emails advertising the study were sent directly to the Product Design Engineering mailing lists for specific year groups at Strathclyde University. The same information was sent to the course leader at Glasgow University / School of Art and passed onwards to students.

- The researchers' personal social and professional network.
- Posters and flyers were placed in public and social spaces in the department of Design, Manufacturing and Engineering Management at Strathclyde and in publicly accessible spaces at Glasgow School of Art.
- Posts were made on Facebook groups, e.g. groups for individual year groups, degree streams departments and department-related social societies.

Recruitment via academic and academic-related social channels provided direct access to the target population.

In Phase 1, participants took part on a volunteer basis. In phases 2 and 3, participants were remunerated for their time. Recruitment in these phases was advertised as part of a group session that included experiments run by other doctoral students. Participants were remunerated £17 for taking part in a 2hour session comprising one of the experiments reported in this thesis, and an unrelated experiment conducted by a different researcher.

In phase 1, participants were placed into the similarity or combination groups depending on their background and experience as at this point the demographic requirements were different for the two tasks. In phase 2, only a similarity experiment was carried out. In phase 3, participants were arbitrarily assigned to the similarity or combination group as at this point the demographic requirements were the same for both experiments.

5.1.2 Stimuli

In all experiments, participants were presented with pairs of design concepts from one of three stimuli sets. A new set of stimuli was created in each of the three phases of research, but experiments that occurred within the same phase utilised the same stimulus sets as inputs. The stimuli set were created via different methods, summarised in Table 5-3 and detailed in Chapter 6.

Anatomy of a stimulus. A single stimulus contained a pair of design concepts presented side by side with a design brief written above (Figure 5-1). Each design concept comprises a sketch with annotations and a typed description below.

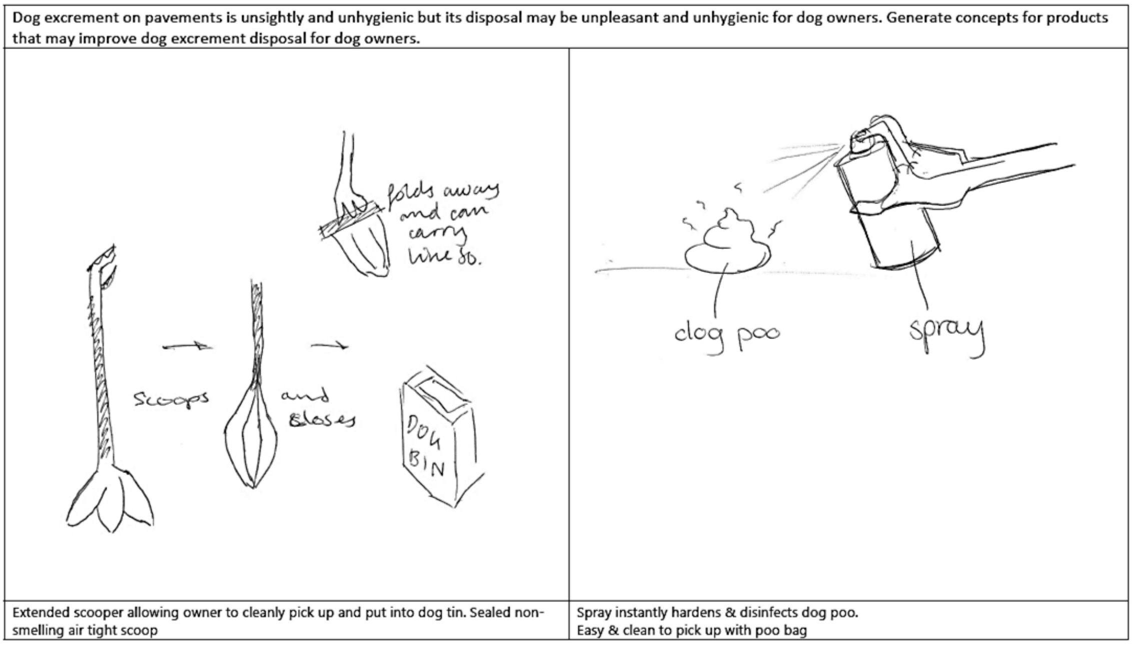


Figure 5-1 – Example of a single stimulus taken from Stimuli Set A.

Source of initial design concepts. The individual concepts that comprise each stimulus were taken from an independent concept generation experiment to which the researcher made a significant contribution. The procedure for this experiment is presented in Appendix 3. The design concepts were created in response to one of four design briefs, listed in Table 5-2. Whenever pairs of concepts were presented alongside a design task, that task was always the one that the design concepts were originally created in response to.

Table 5-2 - List of the four design tasks used in the experiments

Task #	Description
DT03	Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.
DT06	Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.
DT07	Long-distance water transportation may be necessary in drought-prone developing nations but can be problematic due to a lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.
DT09	Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.

Summary of stimuli sets. The three stimulus sets that were used are summarised in Table 5-3. In each set, design concepts are only paired with design concepts from the same task. Within a stimulus set the order within each pair, i.e. whether concepts were presented on the left or the right, did not change. The following details match the columns of Table 5-3.

- Mental representation – the presumed nature of mental design concept representation in the respective stage of research.
- Warmup or main set – shows the number of pairs used in either the warmup or main set of concepts and the design task that they were associated with. In phases 1 and 2 there were no warmups, this was introduced in phase 3.
- Intended manipulation – Sets A and B were created by manipulating the alignability of the pairs to provide *a priori* control of similarity. Set C was created with the intention of the pairs spanning a range of rated similarity.
- Validation method – the similarity manipulation was evaluated using the similarity ratings elicited in each phase.
- Validation outcome - the manipulation for set A failed validation, prompting a new manipulation method in Phase 2. The manipulation for set B failed its intended validation (5 levels of alignability leading to five levels of similarity) but the pairs were shown to span a range of rated similarity which was sufficient for use in Sim-P2. Stimuli set C was then validated against the criteria that the concepts span a range of rated similarity.
- Utilised properties – the properties of the stimuli set, post-validation, that were used in the similarity and combination experiments. In the case of set A, no

properties were utilised because it failed evaluation. In the case of set B, the utilised properties differed from the intended manipulation.

Table 5-3 - Summary of the three stimulus sets used in the three phases of research.

Set	**Experimen ts**	**Mental representation**	**Warmu p details**	**Main set details**	**Intended manipulation**	**Validation method**	**Validation outcome**	**Utilised properties**
A	Sim-P1 Combo-P1	Feature lists	N/A	16 pairs 4 from each of DT03, 06, 07 and 09	Pairs of 'high' and 'low' similarity concepts 2xhigh and 2xlow from each of the four DTs	Similarity ratings from 6 participants	Manipulation not validated	N/A
B	Sim-P2	Features lists/category location	N/A	40 pairs 10 from each of DT03, 06, 07 and 09	Concepts of five levels of similarity based on category position	Evaluation of alignability levels and within-level consistency. Similarity ratings from 11 participants in Sim–P2	Five levels – failed validation. Confirmation of concepts spanning a range of similarity	Concept pairs span a range of mean rated similarity
C	Sim-P3 and Combo-P3	Structured representation	10 pairs All from DT09	30 pairs 10 from each of DT03, 06 and 07	Concepts of varying similarity	Similarity ratings from 37 participants in Sim-P3	Confirmation of concepts spanning a range of similarity	Concept pairs span a range of mean rated similarity

Note: DT = design tasks

5.1.3 General materials

In all experiments, participants were asked to read the general information sheet (Appendix 5A) and then read and sign an experiment consent form (Appendix 5B) if they agreed to take part. A demographic information form was collected and task instructions were presented on paper or using the survey presentation software.

General equipment and environment

- *Sketching and writing implements.* Participants in the combination experiment were provided with HB pencils to sketch with and an eraser. They were also informed that they could bring their own pens or pencils if they preferred. To maintain some degree of homogeneity in the visual properties of the sketches they were not allowed to use rendering markers or any other stationery.
- *Digital presentation software.* During the three similarity experiments, the online survey software Qualtrics software (Version 2018, Copyright © 2022) was used to display images and record similarity ratings, explanations and commonalities and differences from the participants.
- *Standard experiment environment.* The experiments took place in the department of Design, Manufacturing and Engineering Management at the University of Strathclyde. Combination experiments took place in one of two private rooms in the department's research suite, at large 8 or 14-seat desks. Similarity experiments took place in a corner of the open-plan area of the research suite using computers with a minimum of approximately four meters between the participant and any other occupant of the office.

Recurring, task-specific materials. Stimuli were either presented on landscape A4 paper stapled in the top left corner or using the digital presentation software. Participants in the combination experiments were asked to sketch their combined concepts on a sketching template. The sketch template, shown in Figure 5-2, was printed on A4 paper in landscape orientation.

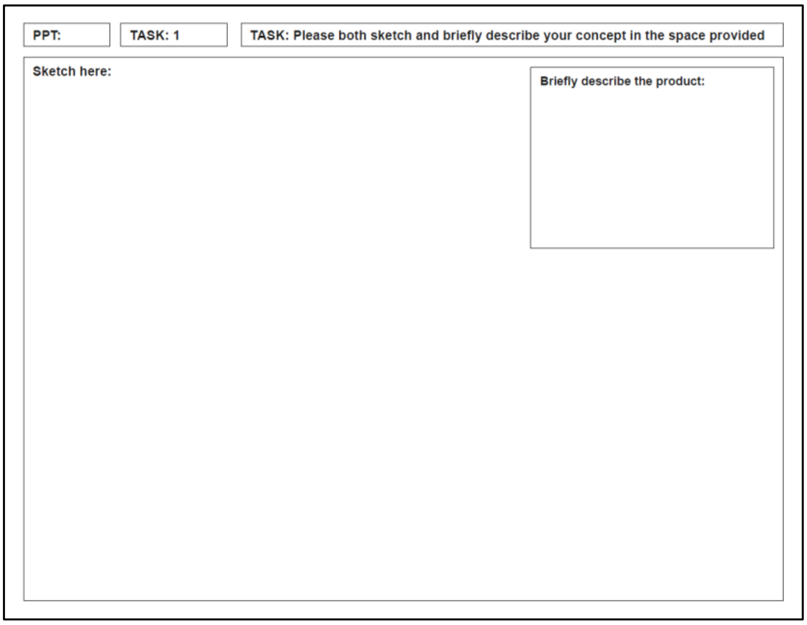

PPT:
TASK: 1
TASK: Please both sketch and briefly describe your concept in the space provided
Sketch here:
Briefly describe the product:

Figure 5-2 – Example of the A4 sketch sheets that the participants use to externalise the combined design concepts.

5.1.4 Experiment tasks

Table 5-4 lists the research methods and experiment tasks used in the similarity and combination experiments in each phase.

Table 5-4 - Summary of the stimuli set, methods, tasks and population used in each experiment.

		Similarity (Sim-Px)			Combination (Combo-Px)		
P	**Set**	**Method / task**	**Pop**	**#**	**Method / task**	**Pop**	**#**
1	A	• Similarity rating task • Observation	E/D familiar	6	• Design concept combination task • Observation	PDE competent	6
2	B	• Similarity rating task • Similarity explanation task • Commonality & difference listing task • Semi-structured interview	E/D familiar	11	-----------------------	------------	---
3	C	• Similarity rating task • Commonality & difference listing task	PDE competent	37	• Design concept combination task • Difficulty rating	PDE competent	30

P = phase of research, Pop = population, # = number of participants. There was no combination experiment in research phase 2.

Four conceptual constraints influenced the implementation of the experiment tasks.

(i) To avoid any influence of stimulus or task familiarity on task performance, participants could not take part in more than one experiment using the same stimuli or experimental task.

(ii) To avoid unintentional priming or biasing of one cognitive process over another, participants could not take part in the commonality and difference listing or similarity rating task if they participate in the combination task. For example, one of the research questions associated with Obj. 3 is to determine whether design concept combination involves a process of comparison or scenario creation. If participants variably used both processes, then prior elicitation of the comparison process (as induced by the listing task and possibly the similarity rating task) may prime the participants to combine concepts via comparison, thereby obscuring any scenario creation processing. This constraint combined with the previous one means that participants could not participate in more than one experiment of any kind.

(iii) In the commonality and difference listing task, participants cannot list both the commonalities *and* the differences for the same pair of concepts. This is because an outcome measure of the listing task is the number of alignable and nonalignable differences listed by participants, but according to the Structural Alignment model (Section 3.2.1.2) noticing commonalities leads to the noticing of alignable differences. Thus, prior elicitation of commonalities and differences by the same participant would result in a skewed increase in alignable differences.

(iv) When similarity rating is carried out in conjunction with the commonality & difference listing task, similarity ratings must be elicited *before* the listing of commonalities or differences. This is because instructing a participant to list commonalities and differences inherently forces a comparison process which could prime subsequent similarity ratings to be conducted via comparison, even if another process such as thematic similarity may otherwise occur.

These constraints were considered when designing the experiments reported in the rest of this chapter.

5.1.5 Selection of statistical tests

Throughout the thesis, statistical tests are used to draw inferences about the populations of interest from the sampled participants. Selection of which statistical test to use depends on a number of factors. Where possible, the specific tests used were the same as those in the prior psychological studies upon which the research design is based. The selection of tests was also informed by the properties of the data, i.e., whether the variables were: normally distributed or not, matched (from the same participant) or not, and whether the assumptions of a given statistical test were satisfied. Based on these factors, statistics tests were selected with reference to the Laerd Statistics statistical tutorials and software guides (Laerd Statistics, 2015a).

In some cases, the selection of statistical tests is informed by findings from prior research (e.g., Section 7.2.1.1) or by specific properties of the data. In such cases, rationale is for the choice of statistical test is presented in the relevant section. For example, rationale is provided in Section 7.2.4.4 for the use of the Wilcoxon signed rank test of group differences (rather than a dependent t-test), and extra steps taken to ensure that the test was robust to the removal of outliers that violate an assumption of symmetry.

5.2 Similarity experiments

5.2.1 Similarity experiment 1

The first similarity experiment (Sim-P1) was carried out in Phase 1. The participants completed the 'similarity rating task' and were observed by the experimenter.

Similarity rating task. The similarity rating task elicits numerical ratings of the similarity of pairs of design concepts. It was adopted from similarity research in the cognitive psychology literature (Markman and Gentner, 1993b, 1996; Tversky, 1977) and used in Sim-P1, Sim-P2 and Sim-P3. Participants were asked to 'rate the similarity of the two concepts' but are not given a definition of similarity. They were instructed to judge the similarity of the artefact represented by the sketches and text, rather than the properties of the sketch itself. Numerical ratings were used to evaluate the similarity manipulation for Stimuli Set A (Section 6.2). Observation of the participants informed procedural changes that are implemented in Sim-P2.

Participants. 6 participants from the 'Engineering/design familiar' population.

Stimuli. Stimuli set A was used for this experiment; 16 pairs of design concepts (2 pairs x 8 design tasks), 1 high and 1 low similarity pair from each task. The manipulation was coding similarity (Jacquard's coefficient), detailed in (Section 6.2).

Materials. The stimuli and rating scales were presented on a computer.

Procedure. The survey guided the participants through the similarity rating task on-screen. The first screen presented them with instructions for the experiment, the second presented the *demographic information form* and the subsequent screens presented the concepts with the rating scales.

One pair of concepts was presented at a time on a computer screen. Each pair was displayed above a Likert-type scale ranging from 1-7, where 1 = "Extremely similar" and 7 = "Not at all similar". The participants rated the similarity of one pair at a time and there was no time restriction. Participants manually clicked an on-screen button to move to the next pair and were not able to move back to prior pairs. Stimuli appeared in a random order using a randomisation feature in the Qualtrics survey software. The experiment took approximately one hour (*M = 596.67s, SD = 214.58s, n = 6).*

The experimenter sat in the eye-line of the computer screen and observed the participants. When the last pair of concepts had been rated the participants were notified on the screen that the experiment was complete and that they should notify the experimenter.

5.2.2 Similarity experiment 2

The second similarity experiment (Sim-P2) was carried out in Phase 2. As per Phase 1, the participants were observed while they carried out the task so that the experimenter could identify any issues with the newly added explanation portion of the task.

- **The similarity rating task**. The similarity rating was the same as described in Sim-P1. Participants provided similarity ratings on a 1-9 scale with 1 being the least similar and 9 being the most similar.
- **The similarity explanation task**. The task was adopted from previous research in cognitive psychology using word pairs as stimuli (Wisniewski and Bassok, 1999). Explanations for similarity ratings can reveal whether the individual based their judgement on the features of two concepts (indicating comparison processing) or a thematic relation between the two concepts (indicating integration processing). It is assumed that the act of explanation does not interfere with the similarity rating process. This is based on prior research, in which similarity ratings of the same stimuli set in a between-groups experiment design have been shown to be "nearly identical" (Wisniewski and Bassok, 1999, p. 217) when participants do and do not provide explanations for their ratings. Explanations were coded to look for evidence of featural or thematic processing. This contributes to Obj. 3.

- **The commonality and difference listing task**. The commonality and difference listing task, also adopted from existing similarity research (Markman, 1996; Markman and Gentner, 1993b), instructed participants to list as many commonalities or differences as they can for a pair of concepts. This was intended to be used as a preliminary test of the algorithmic-level similarity model processes. Owing to an error in assigning participants to list the commonalities or differences for specific stimuli this was not possible, and so the task acts as a procedural pilot for the subsequent listing task in Similarity Experiment 3 thereby contributing indirectly to Obj. 3.
- **Semi-structured interviews**. Semi-structured interviews were conducted after Sim-P2 to get feedback about procedural and conceptual aspects of the experiments. Procedural questions related to e.g., data entry and the duration of the experiment. Conceptual questions elicited introspection from the participants about their thought processes. Participants were informed before beginning the experiment that they would be invited to take part in the follow-up interview after completion of the experiment and that their participation would be valuable, but they were reminded that they were free to leave the experiment at any time.

Five procedural changes for Sim-P2 were borne out of the feedback and observation from Sim-P1. These issues are stated alongside their implications in Table A7-1.

1. The number of total trials was increased.
2. The number of stimuli was increased, and the number of design tasks was decreased thereby reducing the variety of design tasks present.
3. Stimuli were presented on paper rather than on a screen,
4. A requirement was introduced that all stimuli have textual descriptions that explain the content of the sketch.
5. The design task was presented alongside the design concepts.

The unstructured feedback proved valuable but may not have captured the full breadth or depth of insights that were available from the participants. Hence, the semi-structured interviews were added to Sim-P2.

5.2.2.1 Similarity rating, explanation, and listing tasks

Participants. The participants belonged to the ‘engineering/design familiar’ population. Specifically, they were 11 individuals, eight with at least three years of formal product design engineering education, two with master’s degrees in mechanical engineering and 1

PhD researcher with two years of experience in product design research deemed equivalent to the required experience for this population. Participants' demographics are listed in Table A3-2.

Stimuli. Stimuli set B was used for this experiment (Table 5-3). 40 stimuli, ten from each of DTs 03, 06, 07 and 09. The intended similarity manipulation was five levels of alignability based on category position and coding similarity.

Materials. Participants were given the *standard information materials,* a *stimuli booklet* and an additional instruction sheet for the listing task. Stimuli were presented on paper; similarity ratings, explanations and listing responses were done on the computer using a mouse and keyboard. Each participant received a unique booklet with the stimuli presented in a random order. The listing task instruction sheet indicated whether they should list the commonalities or differences for the corresponding pair of design concepts in the stimulus booklet. Participants used the same booklet for the similarity rating and the listing tasks and thus saw the concepts in the same order for each task. The experiment took place in the *standard experiment environment*.

Procedure. The experiment was carried out in two parts: (i) similarity ratings and explanations, and (ii) commonality and difference listing. Participants were given full instructions for the similarity rating and explanation tasks at the beginning of the experiment and were told that they would complete another task in part 2. They were told that the task would involve making more judgements about the design concepts and that full instructions would be provided in part 2. The nature of the commonality and difference listing task was obscured to mitigate the risk that knowledge of an upcoming comparison task would force featural processing and thus inhibit any thematic processing that would otherwise take place.

Participants provided a similarity rating and explanation for each of the 40 pairs of concepts. The similarity scale in the first part of the experiment ranged from 1 (not at all similar) to 9 (extremely similar), with 5 being 'moderately similar'. They were presented with the rating scale and explanation text box simultaneously but were instructed to provide the numerical rating first and then type an explanation without changing their initial rating. After providing a rating and explanation the participant could move to the next page at their own pace, but once they had moved on, they could not return to a previous pair. No time limit was placed on this part of the task.

Participants were then prompted to return to the start of the booklet to begin the commonality and difference listing task. They were given 1 minute to respond to each pair of concepts and they were timed by the experimenter who asked them to move on after 1 minute.

Explanation coding. The coding procedure is shown in Figure 5-3. The participants' explanations were coded as 'featural' (F), 'thematic' (T), or 'other' (O) based on definitions from the literature.

- **Explanations were to be coded as 'feature-based' if they contained commonalities or differences.** For example, *"Both have [the] exact same function - tracks and displays food date, but one has [the] additional function of keeping food fresh for longer"* the participant specifies the function of the concepts as a common feature and a secondary function as a difference.
- **Explanations were to be coded as 'thematic' if there was any mention of an external and complementary relation between the two concepts.** For example, *"One concept collects water from the air, the water is then transferred to the other concept which transports it".* In this hypothetical example, the transfer of the water from one concept to the other is a thematic relation and both concepts are used together in complementary roles in a scenario.
- **Any response that could not be coded as one of the above was coded as 'other'.**

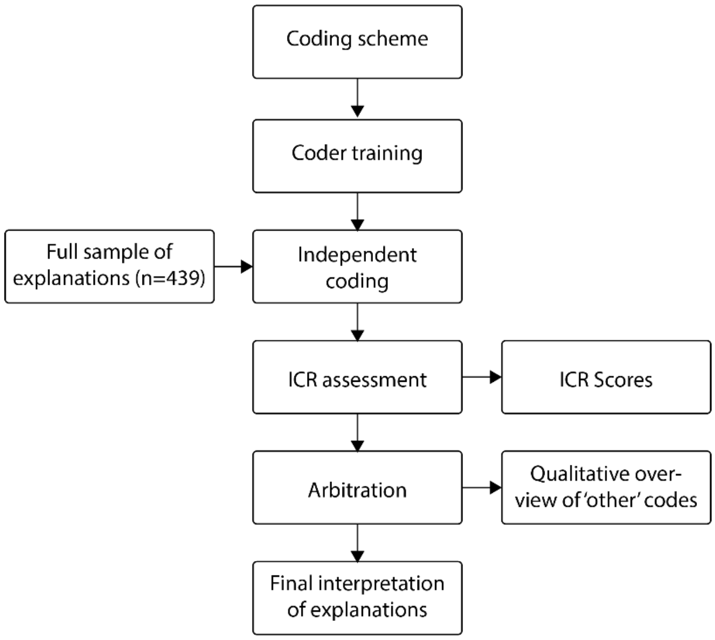


Figure 5-3 - Coding procedure for similarity explanations in Similarity Experiment 2 (Sim-P2).

The researcher completed the first pass of coding using the definitions given above on the full sample of 439 responses. A second rater was then trained on the same definitions using a coding handbook (Appendix 5C) before rating all responses. The contingency table for the two judges is shown in Table 5-5. The initial agreement was 94.99%[13]. 22 responses were arbitrated by a discussion leading to the final interpretation of explanations.

Table 5-5 - Coding contingency matrix for similarity explanations

		Rater 1 (Researcher)		
		F	T	O
Rater 2 (independent)	F	410	0	9
	T	0	0	1
	O	12	0	7

[13] In this instance a simple percentage agreement was deemed sufficient for describing agreement between raters. This percentage agreement indicates high agreement between judges, but corresponds to a Krippendorff's alpha value of 0.41 which indicates poor agreement. This can happen when most of the data fall into a single category (Zec *et al.*, 2017).

439 explanations coded, grey boxes indicate agreement (n=417)

5.2.2.2 Semi-structured follow-up interview

Following the completion of the similarity experiment, participants were asked if they would like to participate in a semi-structured follow-up interview to help evaluate the experimental procedure and to gain insights into their cognitive processes through introspection. Eight open-ended questions were asked of seven participants who volunteered to take part.

Method. Each interview was audio recorded and transcribed.

List of questions. Eight questions were asked.

1. Do you have any general thoughts on the experiment; any comments you would like to make?
2. Were there any elements of the experiment which you found to be particularly difficult?
3. How do you feel about the means of data entry (i.e. via laptop)
4. Do you feel that the accuracy of your similarity judgements was consistent throughout the experiment, or do you think it changed?
5. Do you feel that the manner in which you determined similarity changed throughout the study?
6. Do you feel that you had enough time to list the commonalities or differences of a pair of concepts?
7. Please comment on your energy/effort levels towards the end of the experiment.
8. Do you have any suggestions for how the experiment might be improved?

Analysis. The responses to the questions were interpreted by the researcher. Through iterative rounds of analysis, the content of the responses was interpreted, issues were identified, and their implications and potential resolutions were logged. These data are presented in Table A7-3.

5.2.3 Similarity experiment 3

The third similarity experiment (Sim-P3) was carried out in Phase 3. Both the **similarity rating task** and **commonality and difference listing task** were used to test the hypotheses associated with the proposed model of design concept similarity judgements.

Five procedural changes were implemented based on the responses to the semi-structured interview questions in Sim-P2.

1. A practise trial was introduced
2. Verbal assurances were given to participants that they need not worry about repetition in their answers, but they were also instructed not to copy and paste their responses
3. Some stimuli were changed (see Section 6.3.4.1)
4. The time constraint on the commonality and difference listing task was removed
5. Some hardware was swapped owing to issues with the keyboards used

5.2.3.1 Similarity rating and commonality and difference listing tasks

Participants. 37 participants were drawn from the 'PDE competent' population.

Stimuli. Stimuli set C was used in this experiment. This comprised two stimuli booklets, one for a warmup (10 pairs x 1 task) and one for the main rating task (10 pairs x 3 tasks). Each participant saw pairs spanning a range of similarities in one of ten random orders.

Materials. The materials were the same as those used in Sim-P2.

Procedure. The experiment was carried out in three sequential parts: (1) warmup similarity rating exercise, (2) main similarity ratings, and (3) commonality or difference listing. At the beginning of the experiment, the participants were given full instructions for the similarity rating and explanations task and were told that they would complete another listing task in part 2. Full instructions for part 2 were given after the completion of part 1. The experiment took place in the standard experiment environment. No time limit was placed on the tasks.

Participants were instructed that the purpose of the warmup was to let them become familiar with the rating scales and that their responses would not be analysed. Participants rated the similarity for the 10 warmup pairs from 1 (not at all similar) to 9 (extremely similar), with 5 being 'moderately similar'. After providing a rating the participant could move to the next page at their own pace but could not then move back to a previous pair. Upon completion, they were asked if they had any questions before moving on to part 2 where they rated the similarity for the 30 pairs using the same procedure as before.

Once the participants had completed the ratings they were instructed to return to the start of the booklet to begin the commonality and difference listing task. Unlike Sim-P2, no time limit was placed on the listing task; they were instructed to continue to list commonalities or differences until they felt they had run out.

5.2.3.2 Difference coding

Differences were coded as alignable or nonalignable using a deductive coding approach, meaning that the definitions of alignable and nonalignable differences were adopted from the psychology literature and applied in a design context. The coding process was carried out in two rounds, both of which involve the same core elements (Figure 5-4). That is, codes were defined, a sample of differences was coded by the researcher, the same sample was coded by an additional judge or judges and disagreements were arbitrated. At the end of round 1, conceptual issues were identified in the application of the scheme leading to refined code definitions for round 2. At the end of round 2, a satisfactory level of reliability was achieved.

As per Markman and Gentner (1993, 1996) a conservative measure of alignable differences was used. This means that participants may have noticed an alignable difference but failed to report it as such. This should make relevant hypotheses more difficult to support. For example, a participant tasked with listing the commonalities and differences between a car and a motorcycle may have stated: “Cars have four wheels, and motorcycles don’t”, i.e., listing a nonalignable difference. It is possible that during the experiment the participant was thinking “Cars have four wheels, and motorcycles have two”, i.e., listing an alignable difference. Judges were encouraged to make their decision based on the content of the response, and not their inference about what the participant may ‘actually’ have been thinking.

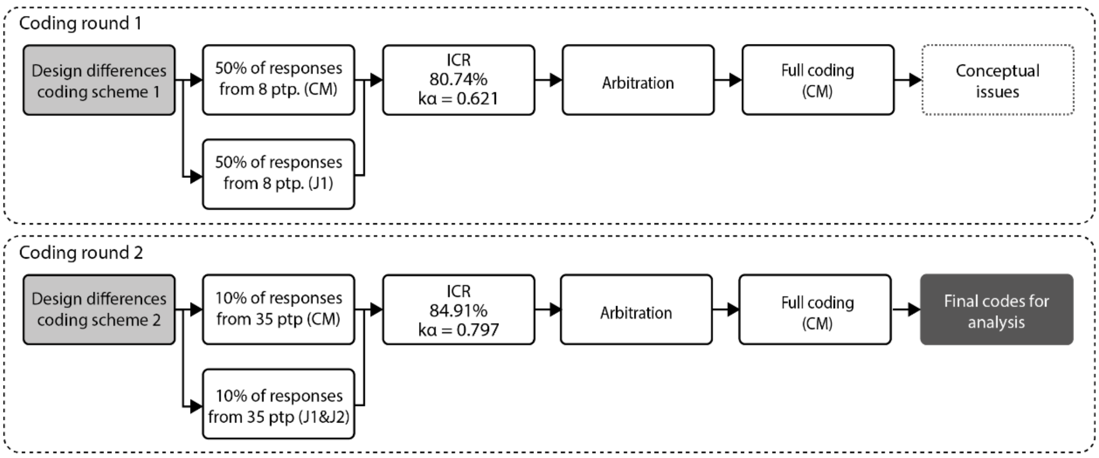


Figure 5-4 – Overview of the coding process for alignable and nonalignable differences, Sim-P3

Round 1. A coding book was created using definitions and examples of alignable and nonalignable differences from studies in the psychology literature (Appendix 5C). The data from the first 8 participants to complete Sim-P3 was cloned to conduct an initial coding trial. 32.5% of the responses from the first 8 participants were selected using a random number generator and coded. Of 135 responses, 109 were the same and 26 were different, giving a percentage agreement of 80.74%. Krippendorff's Alpha (kalpha) = 0.621. To identify the source of disagreement the 26 differences were arbitrated. The results of the arbitration, combined with the subsequent application of the coding scheme to the full sample of data, revealed issues with the coding scheme that were addressed in round 2.

Round 2. The coding scheme was the same as described previously, with the addition of an 'other' code and five new clarifying rules to help resolve ambiguities in the data (Appendix 5D).

A sample of 245 differences was created for a test of inter-rater reliability (IRR). This is 13.45% of the 1822 differences listed by 35 participants in response to the commonality and difference listing task (two participants were excluded from analysis, see Section 7.2.2.2). This sample was derived by randomly selecting responses to the listing task (containing an average of 4.08 differences each). This comprises responses from 27/35 participants to 26/30 stimuli. Of the 245 responses, there were 13 cases where at least 2 of 3 judges determined that the participant has spread a response (AD) over two lines. These cases were collapsed so that they would not be counted twice and artificially inflate the reliability score. Of the 232 differences, the judges agreed on 197 (84.91%) and disagreed on 35 (15.09%), kalpha = 0.797. This result was taken as sufficient for moving on to coding the full coding scheme. The researcher reviewed the disagreements to gain an understanding of the causes of the differences between the judges and then coded the entire dataset.

5.3 Combination experiments – specific materials and methods

This section reports the two combination experiments conducted in Phases 1 and 3, termed Combo-P1 and Combo-P3 respectively. Table 5-4 shows the methods used in each combination experiment.

5.3.1 Combination experiment 1

This section reports the experimental methods for Combo-P1. Participants were observed while they carried out the design concept combination task. In the design concept combination task, participants were presented with pairs of design concepts accompanied by the brief that they were initially generated in response to, and were asked to combine them to create a new concept that addressed the same brief. The participants were not given a definition of 'combination' as the task was designed to elicit combination processing from the participants in whatever form the participants understood the term to mean. This task acted as an initial pilot study for the design concept combination paradigm used again in Combo-P3, thereby contributing indirectly to Obj. 4. It also provided data used in an initial trial of concept 'type' coding, contributing to the requirement for coding schemes. Observation of the participants informed procedural changes that are implemented in Combo-P3. Participants were also timed to determine an appropriate duration for the next combination experiment (Combo-P3).

Participants. 6 people participated; all participants had first degrees in PDE or a related discipline and were PhD students at the time of participation. 5 participants had undergraduate and/or Master's degrees in product design engineering. 1 participant had an undergraduate degree in sports engineering but was included based on their research and teaching experience in PDE. Participants' demographics are listed in Table A3-3.

Stimuli. Stimuli set A was used for this experiment; 16 pairs of design concepts, 2 pairs from 8 design tasks, and 1 similar and 1 dissimilar pair for each task. The manipulation was coding similarity (Jacquard's coefficient), detailed in Section 6.2.

Materials. Participants were given the *standard information materials,* a *stimuli booklet, sketch sheets,* and three HB pencils. All information presentation and responses were done on paper. The stimuli were of the same format as described in Section 5.1.2 and included the design brief. Each participant received a unique booklet that included the same stimuli in a random order

Procedure. Participants were instructed, for each pair, to combine the two concepts to create one new concept that addressed the same brief. No time limit was placed on each sketching trial. Specific instructions that elicit combination are as follows:

> *"For each task, you will be presented with a problem brief, two existing concepts and (where applicable), descriptions of those concepts. Please read the problem brief and*

then sketch a new concept that is a combination of the two existing concepts. Sketch this new concept on the sheets given to you; each sheet has a number which corresponds to the tasks in the stimuli packet.

Your sketch of your concept need not be highly detailed and need not include details such as dimensions. You should try to convey the key components of the concept, and you may annotate your sketches."

The experimenter sat in the eye-line of the participant's stimulus book and sketch sheets and observed and timed the participants using a digital timer. When the last pair of concepts had been rated the participants were notified that the experiment was complete and that they should notify the experimenter.

5.3.2 Combination experiment 2

This experiment (Combo-P3) was carried out in Phase 3 and utilises Stimuli Set C. The data collected from the **design concept combination task** were used in the development of the design concept combination coding scheme (Section 5.3.2.2), thereby contributing to Obj. 4 . The combined concepts were analysed (Chapter 8) to address Obj. 4*.* **The difficulty rating task** provided subjective ratings of how difficult it was to combine each pair of concepts. After combining a pair of concepts to create a new one, participants were asked to rate how difficult it was to combine that pair. Unlike Combo-P1 participants were not observed.

5.3.2.1 Tasks

Participants. 30 people from the 'PDE competent' population participated. None of the participants in this experiment took part in any of the other experiments presented in this thesis.

Stimuli. The main set of 30 pairs of design concepts from Stimuli set C was used for this experiment. Note that the warmup booklet used in Sim-P3 was not used in Combo-P3.

Materials. The materials were the same as those for Combo-P1.

Procedure. Participants were given three minutes to complete the combination task and sketch their concept. At the end of the three minutes, they were told to finish what they were doing and rate how difficult it was to combine the concepts. They were informed that if they were finished before the 3 minutes, they could verbally inform the experimenter and move on to the next pair.

The instructions given to participants were based on those in Combo-P2 but with minor changes to the paragraph formatting and language. Participants were asked to combine concepts to create a new concept that addresses the same brief.

> *"1. Combination task instructions*
>
> *For each task you will be presented with a problem brief, two existing concepts and (where applicable), descriptions of those concepts.*
>
> *Your task is to combine these concepts into a new concept that satisfies the brief for that task.*
>
> *Sketch this new concept on the sheets given to you. You will be given 3 minutes to sketch a new concept, however you may move on before this time if you are satisfied with your concept."*
>
> *2. Difficulty rating*
>
> *After each sketch please rate how difficult it was to combine the two concepts on the sheet provided."*

The experimenter sat in the same room as the participant facing perpendicular to the orientation of the participant, approximately two meters away. The setup still allowed the experimenter to inform the participant when 3 minutes had passed, make sure they completed the difficulty rating, and receive verbal acknowledgements if the participant moved on to a new task early. When the last pair of concepts had been rated the participants were notified by the experimenter that the experiment was complete and that they were free to leave.

5.3.2.2 Combination type coding

The coding of combination types took place over three rounds (Figure 5-5). The first round began with combination-type codes from the Dual-Process model of conceptual combination that were iteratively evolved to be applicable for coding design concepts. The second round was an evaluation of this coding scheme that resulted in poor inter-rater reliability. This scheme was modified again and evaluated in the third and final round of coding, resulting in satisfactory reliability.

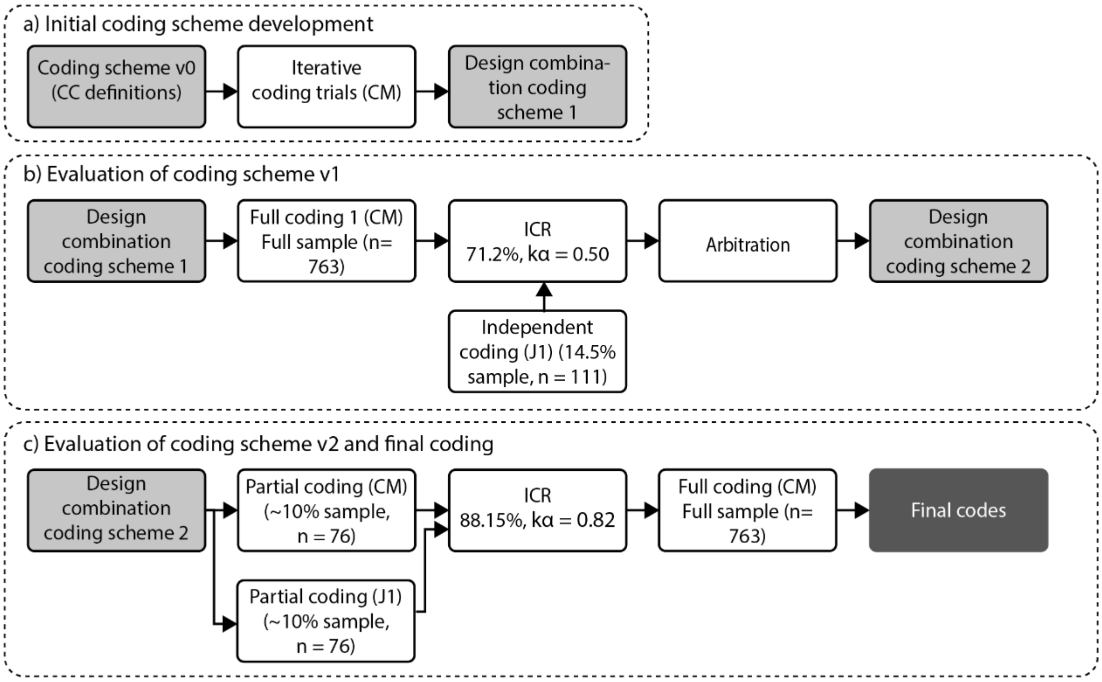


Figure 5-5 – Overview of the coding process for combination types, Combo-P3.

Initial coding scheme development (Figure 5-5a). The initial phase of coding scheme development began with nominal definitions of combination types from the psychology literature (Coding scheme v0) and lead to the development of the first version of the design concept combination coding scheme (Coding scheme v1). For each design concept, an attempt was made to apply one of the initial codes from the conceptual combination literature. If none of the codes were applicable a new code was defined.

Four initial definitions were adopted from Wisniewski (1996), p.438-439.

- Property-mapping - one or more properties of a constituent were asserted of the referent of the combination, as in “*grey* clay” for ELEPHANT CLAY, “*thin* rake” for PENCIL RAKE, and “*pony* with stripes”, for TIGER PONY.
- Hybridisation – the combined concept involves combinations of the two objects or entities that were both of the objects, as in “a very large heavy creature sharing properties of both an elephant and a moose,” for moose elephant, or “a combination ladder/rake,” for ladder rake.
- Relational – the combined concept involves a relation between two objects, as in “box that *holds* ladders” for LADDER BOX, “squirrel that *chases* cars,” for CAR SQUIRREL, and “robin that *eats* snakes,” for SNAKE ROBIN.
- Other – the combined concept cannot be coded as any of the above.

The outcome of the initial coding scheme development is shown in Table A6-1. The three forms of combination from the Dual-Process model of conceptual combination (Wisniewski, 1997a) evolved into a scheme comprising 6 types of response to the design concept combination task. These codes were Featural (F), Relational (A), Non-combinational (NC), Insufficient information to interpret sketch (I) and Unclear how concept addresses brief (U). The featural and relational codes correspond to the property-mapping and relational combinations from the initial definitions in the bullet list above. Unexpectedly, in many responses participants had successfully created a design concept but there was no evidence that combination had taken place. For some codes, multiple sub-types were identified that were deemed to be unnecessarily granular but were included in the descriptions of each code to help the coders.

Coding round 1. Coding round 1 was carried out to evaluate and improve 'Combination Coding Scheme v1' by attaining a measure of inter-rater reliability and, if necessary, adapting the coding scheme. A coding booklet was created that contained a coding scheme and coding procedure (Appendix 6C).

The researcher coded the full sample of responses which consisted of 763 design concepts created by 29 participants in Combo-P3. This initial coding of the entire sample was conducted to test the applicability of Coding scheme v1, as up until this point the scheme had not been applied to all design concepts. It was possible to apply the coding scheme without any substantive issues.

To gain a measure of inter-rater reliability, an independent judge (J1) coded a randomly selected 14.5% of the same sample. This subsample excluded the participant that was removed from the analysis (Section 8.1.2.1). It also did not include any NR codes (no response) to avoid artificially inflating the reliability score with responses that were trivially easy to code. The researcher and J1 agreed on codes for 71.2% of the design concepts (79/111), kalpha = 0.4936. This indicates poor reliability between the two judges. Thus, although the scheme could be applied consistently by the researcher, it was not sufficiently robust to allow an independent coder to apply the scheme consistently.

To identify weaknesses in the coding scheme, arbitration was conducted to identify the reasons for the 32 disagreements. The judges each presented their rationale for the code they selected, discussed the disagreement and attempted to arrive at a consensus. Consensus was achieved in all 32 cases and the reason for disagreement was noted.

Of the 32 disagreements, the researcher was correct for 15, J1 was correct for 12, and neither judge was correct in 5 instances. The reasons for disagreement were analysed by the researcher and five types were identified.

- Ambiguous (3/32) – the correct code was 'Ambiguous' but one or both judges were incorrect.
- Difference in interpretation (7/32) – disagreement can be attributed to differing interpretations of the design concept. When the judges discussed their interpretation one judge always deferred to the other's interpretation.
- Insufficient info (3/32) – The judges disagreed on the interpretation but could not agree on a 'correct' interpretation owing to a lack of information in the sketch.
- Issues with the definition of a code (4/32) – There were issues regarding the clarity of the ambiguous (1/4) and relational (3/4) codes.
- Judge mistake (16/32) – disagreement was attributed to improper application of the codes, either by the researcher (5/15), J1 (9/15), or both judges (1/15).

If the 15 disagreements that were attributed to mistakes by the judges were not made, the maximum agreement that could have been achieved was 85.6% (95/111). This was interpreted to mean that the coding scheme still required development.

Two issues were identified that could be addressed to improve the coding scheme. First, the definition of 'relational' combination was defined as comprising a special kind of combination in which 'an entity in one concept has been substituted for an entity in the other design concept'. This kind of combination was a form of ambiguous combination that had been misattributed as relational. These substitution combinations can be construed as either relational or featural depending on the interpreter's frame of reference and should thus be ambiguous. Secondly, the definition of a relational combination was still not clear. This was addressed by adding a sufficient condition and extra information for defining relational combinations. This resulted in the creation of Coding scheme v2 (Table 5-6).

Coding round 2. Coding round 2 was conducted to evaluate Coding scheme v2 and code the full set of data to produce the final codes used in the analyses. To determine whether the new coding scheme could be applied with greater interrater reliability, a new, randomly selected 10% of the data were selected and coded by the researcher and J1 (n=76/763). The judges agreed on 67/76 codes (88.16%), kalpha = 0.816. Following the heuristics established by (Krippendorff, 2018, p.241-242) this was taken as an acceptable level of reliability. Subsequently, the entire dataset was coded by the researcher using

Coding scheme v2. This produced the final set of codes that were used in the analysis of Combo-P3 (Chapter 8).

Table 5-6 – Combination coding scheme v2

Code	**Label**	**Code description**
Successful trial - Combination		
Featural	F	The newly created, single design artefact contains unique elements from the artefacts in both base concepts.
Relational	R	The combined concept has been created by relating two entities together. The two entities must be from different base concepts and both must be novel artefacts or specific existing products designed for use in a system. The original entities may have been modified (sometimes extensively) to facilitate the creation of a new relation between them, but you should be able to infer that the relevant artefact in the output is a modified version of one of the base artefacts. A sufficient condition for a relational combination is the presence of two spatially distinct entities that can be attributed to entities from different base concepts. Relational combinations can also be spatially co-located entities where it can be inferred that the base concepts have been related by an external structural relation.
Ambiguous	A	Ambiguous combinations are those that can be coded as both featural or relational, depending on one's frame of reference.
Successful trial - Non-combinational ideation		
Non-combination	NC	The new design concept addresses the brief but there is no evidence that combination has occurred. There are at least three kinds of non-combinational concept: • entirely new concepts, • modifications of one base concept (without elements of the other), • artefacts that builds on commonalities of both base concepts (but without any transfer of elements that would indicate a featural combination).
Unsuccessful trial		
Other	I	Insufficient info - The representation of the design concept cannot be interpreted.
	U	Unclear how address brief - A new design concept has been created. It is possible to infer the intent of the designer but the concept does not appear to address the design brief.
	O	Other – the concept cannot be coded by any of the other codes

5.4 Summary

This chapter has described the experimental methods for all of the studies in this thesis. This includes general materials and methods used throughout the studies (Section 5.1), three similarity experiments (Sim-P1, P2 and P3, Section 5.2) and two combination experiments (Combo-P1 and P3, Section 5.3). This chapter is referenced in Chapters 6 (stimuli creation), 7 (similarity results) and 8 (combination results). The correspondence between these chapters can be seen in Figure 4-6.

6 STIMULI CREATION

As established in Chapter 4, the first requirement for empirical research (R1) was to create pairs of design concepts of varying similarity for use in the similarity and combination experiments. This chapter reports three phases of stimuli creation, which can be seen in context on the research map in Figure 4-6. In total, three sets of stimuli were created using three different methods, one for each phase of the research. Each stimuli set was created to facilitate the similarity and combination experiments in that phase. The sets were changed between phases to improve on the prior phase by implementing a different manipulation or by introducing additional control measures to increase the homogeneity of the stimuli within the set.

For each set, evaluation criteria were defined and the sets were evaluated using the similarity ratings elicited in the similarity experiment in the respective phase. All stimuli sets were created to address at least the minimum requirement of varying in similarity (e.g., groups of high and low similarity or a range of similarity ratings). Two sets of stimuli were initially created with the more ambitious goal of directly controlling the similarity of pairs of design concepts by manipulating their alignability. Neither set satisfied these additional criteria and were either abandoned or re-evaluated in terms of similarity alone. The chapter begins with a review of methods for creating pairs of concepts that vary in similarity, drawn from the psychology literature (Section 6.1). Sections 6.2 - 6.4 report the creation of stimuli sets A, B and C (phases 1, 2 and 3 respectively).

6.1 Methods for concept creation

The first requirement for empirical research (R1) was to create pairs of design concepts that varied in similarity. No studies could be found in the design literature in which early-stage, sketch-based design concepts were used as stimuli, nor in which the similarity of pairs of design concepts had been manipulated. To inform the stimuli creation process, inspiration was taken from prior experimental research into similarity judgements and conceptual combination in the psychology literature. Four methods were reviewed along with their benefits and limitations as methods for creating pairs of design concepts that vary in similarity.

6.1.1 Intuition and experimental evaluation.

Pairs of highly similar and dissimilar concepts can be created using a mixture of researcher intuition and independent ratings. In a study of conceptual combination and emergent properties, Wilkenfeld & Ward (2001) first created 90 pairs from various categories (including musical instruments, tools, natural items and 'manufactured items'). 45 pairs were assumed to be similar and 45 were assumed to be dissimilar. 53 independent raters were then asked to rate the similarity of the pairs and these ratings were used to select the 8 most and least similar pairs.

Benefits. In comparison with the other three classes of methods, the use of intuition as a first-pass approximation of similarity requires relatively few resources. The method is not limited to use with any particular kind of stimuli.

Limitations. The need for independent ratings may increase the resource requirements for participant recruitment. The exclusive use of human judgements means does not account for two different types of similarity and may introduce unforeseen biases.

6.1.2 Semantic similarity

Semantic similarity generally refers to some measure of the likeness of the semantic content of words or passages, i.e., the meaning of the text.

- The distance between the relative position of words in semantic networks such as wordnet (Miller, 1995) or the EDR concept dictionary (Yokoi, 1995) (e.g., Nagai and Taura, 2006; Taura, 2016, Chapter 8) can be used as a measure of semantic similarity.

- Similarly, in Latent Semantic Analysis (LSA) (Landauer, 1997) the meaning of words or passages is derived from the contextual usage of the words in a large corpus of text. LSA has been used to assess the similarity of head-nouns in noun-noun pairs in conceptual combination (Xu and Paulson, 2013).
- Latent Dirichlet Allocation (Blei *et al.*, 2003) (a machine learning technique) infers 'topics' from within texts, and those that share dominant topics in similar proportions are judged to be more similar overall. It has been used to estimate the semantic distance between ideas for solutions to environmental and social problems from the openIDEO database (https://www.openideo.com/) (Chan and Schunn, 2015).

Benefits. Lack of reliance on human judges, proving greater repeatability and lowering resource demands.

Limitations. The semantic similarity methods are limited in that they require text-based representations. This problematic for manipulating the similarity of design concepts, since it cannot account for a perception of similarity derived from a sketch of a design concept.

6.1.3 Artificial stimuli and manipulation of elements

The alignability of certain types of pictorial stimuli can be controlled by creating artificial stimuli in which the constituent objects and relations are manipulated manually by the researchers. Assuming that similarity is a product of a structural alignment view of similarity, this facilitates indirect control over similarity. Markman and Gentner (1996) use this technique to create triads of pictorial scenes. Wisniewski and Middleton (2002) created pairs of 'novel microorganisms', comprising geometric shapes represented pictorially and given fictitious names.

Benefits. Grants precise control over alignability.

Drawbacks. The use of artificial design concepts would limit the external validity of any studies, compared with e.g., design concepts created by designers during the design process. The method is also heavily dependent on assumptions about mental design concept representation and what constitutes an object or relation in design. Does not account for the potential role of thematic relations in similarity judgements or combination.

6.1.4 Category location as a proxy for featural commonalities.

The location of concepts in a concept ontology has been used as a proxy for similarity (Markman and Gentner, 1993b; Wisniewski and Markman, 1993). This is because membership in a taxonomic category depends on featural similarity to known members of that category. Taxonomies are systems in which categories (concepts) are related to one another by class-inclusion relationships, with subordinate categories being members of superordinate categories (Rosch *et al.*, 1976). Members of the same superordinate class share many common properties by virtue of their common category membership. For example, armchairs, rocking chairs and reclining chairs are all members of the 'chair' category and all share many features. An example from Markman and Gentner (1993b) can be seen in Figure 6-1. 'Ontological distance' was used as a proxy for similarity, where ontological distance is defined as the number of nodes that have to be traversed to get from one of the lowest level nodes to another.

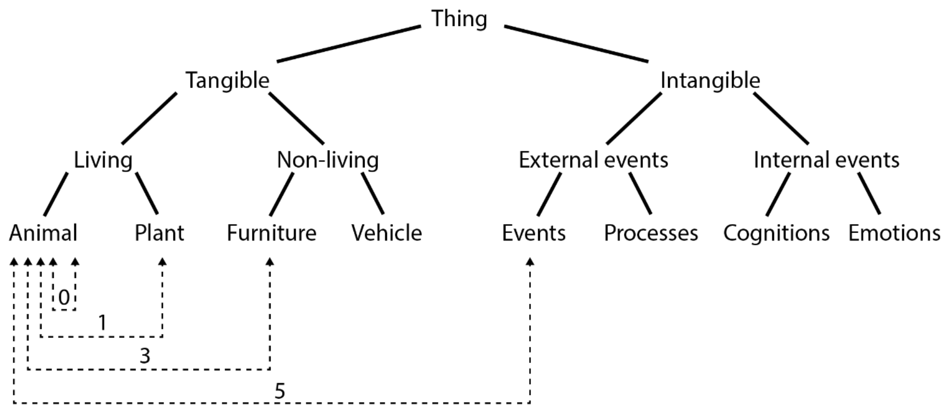


Figure 6-1 – An example of a 'concept ontology' re-drawn from Markman and Gentner, (1993b). Numbers denote ontological distance.

Benefits. Successful creation of a single category would allow for the selection of multiple pairs of concepts at varying degrees of similarity.

Limitations. The method relies on being able to categorise design concepts in the kinds of taxonomic concept ontologies shown in Figure 6-1. At the time of reviewing this method, it was not clear whether design concepts could be classified in taxonomic hierarchies, or what the properties of those classifications would be (but see Section 6.3 for a test and evaluation of this method in a design context).

6.1.5 Summary

Four methods for manipulating the similarity of concepts were reviewed. These informed the stimuli creation methods that were used in each phase of research.

- The manipulation of alignability, as is the case in the artificial stimuli, influenced the coding-similarity method used in Phase 1 (Section 6.2).
- The concept categorisation method was adapted for use in a design context in Phase 2 (Section 6.3)
- The intuition and independent judgement method highlighted the benefits to ecological validity in having human validation of stimuli; an idea that permeated across all three phases.

The measures of semantic similarity were deemed not to be compatible with the early-stage design concepts specified in the research focus (Section 1.1.4.2).

6.2 Phase 1: stimuli set A

Stimuli Set A was created via a novel method of coding design concepts and computing similarity from the common and different features of the pair. The goal was to produce two groups of high-similarity and low-similarity concepts. The method was evaluated by the designer similarity judgements elicited in Sim-P1. The success criteria for the evaluation were that (i) the similarity ratings should be significantly higher in the high similarity condition than in the low similarity condition, and that (ii) the similarity ratings should not differ significantly *within* each of the two groups.

6.2.1 Theoretical basis and concept representation

As part of research phase 1, when Stimuli Set A was created it was assumed that similarity judgements were the result of predominantly featural processing and that thematic processing was negligible. It was further assumed that design concepts can be represented by coding them in terms of their Function, Behaviour and Structure and that by manipulating the degree of common and different F, B and S variables that it would be possible to achieve *a priori* control over human similarity judgements. These assumptions arose from the early stages of work carried out by Dr. Hay with support from the researcher which later evolved into a coding scheme for analysing exploratory design ideation (Hay *et al.*, 2020).

The design concepts were coded with qualitative descriptions of the function, behaviour and structure of the artefact. This coding scheme was an early prototype developed by Dr. Hay and as such the coding scheme itself is not presented as a contribution in this thesis. The coding scheme comprised three types of code [14]:

- **(F) Overall function:** a purpose the product fulfils in relation to the goal of the design problem, where the goal is some future desirable state to be attained.
- **(B) Purposeful behaviour**: a particular aspect of the product's behaviour that allows it to fulfil an overall function.
- **(S) Function carrier**: a particular part of the product's physical structure that is fundamentally involved in producing a purposeful behaviour, and in turn, fulfilling an overall function.

The codes were developed iteratively by examining each concept, defining a code, and grouping similar codes over time. Concepts could be coded with more than one code from each class. An example of a coded concept is shown in Figure 6-2, Figure 6-3 and Table 6-2.

To create pairs of high and low alignability concepts, the Jaccard coefficient of all pairwise combinations of concepts in each design task was calculated. The Jaccard coefficient is a measure of similarity for two sets of data, where the value is a function of the number of matching and mismatching codes. It is a limited form of the Tversky index (Section 3.2.1.1) where $\alpha = \beta = 1$. Concepts were taken from the upper and lower ends of the range of Jaccard coefficient values to create the high and low similarity groups.

The Jaccard coefficient is expressed as:

$$J(x, y) = \frac{|x \cap y|}{|x \cup y|}$$

Or,

$$Jaccard\ coefficient = \frac{the\ intersection\ of\ two\ sets}{the\ union\ of\ two\ sets}$$

[14] These three codes and the exact wording presented here were originally defined by Dr. Laura Hay. They were working definitions at the initial time of creation and should be considered to be temporary working definitions.

6.2.2 Method

Concept sample. Design concepts were taken from an ongoing idea generation experiment (Appendix 3B). At the time, 36 participants had completed the experiment in which they could create up to 3 design concepts in response to 10 design tasks. This led to a sample of 644 design concepts across 10 design tasks. Eight of ten design tasks were used. Two tasks (DT14 and DT19) were excluded because the solutions predominantly involved the creation or modification of urban infrastructure (e.g., bridges, pavements and roads) and as such were deemed to be outwith the PDE domain.

Filtering. To mitigate the chance of confounds stemming from inconsistencies in representation modality or quality, the sample of concepts was filtered to homogenise the quality of representation of the design concepts. Concepts were removed by the researcher if they did not adhere to any of the criteria.

- Sketch - The response contained a sketch of a design artefact (as opposed to only a textual description).
- Internal consistency - The sketch of the artefact matched the annotations and textual description provided.
- Address the brief - It could be inferred with little ambiguity how the presented sketch and annotations could represent a concept that could address the brief.
- Sketch quality – The sketch was deemed to be of relatively high quality as decided subjectively by the researcher.

The filtering process included a degree of subjectivity on the part of the researcher, particularly with regard to the third criterion. Table 6-1 shows the number of concepts in each design task that were coded with the 'hqs' code.

Table 6-1 – Design tasks and the number of associated design concepts used in Phase 1.

	Design Task							
Task #	**DT03**	**DT06**	**DT07**	**DT08**	**DT09**	**DT12**	**DT15**	**DT20**
n_all	86	75	74	73	85	84	80	87
n_hqs	20	20	20	18	32	40	41	45

N_all = number of design concepts associated with each design task when the concept samples were cloned from the behavioural study. n_hqs = the number of those codes coded as a 'high-quality sketch'.

Pair creation. The Jaccard coefficient for all pairs in a design task was calculated using the cluster analysis function in NVivo Version 11. This returns a value between 1 and 0 for all pairs. Pairs were then selected by the researcher from near the top and bottom ends of the ranges of Jaccard coefficient values, excluding values of 1 (identical) or 0 (no common features), including only pairs where both sketches were coded as high-quality sketches.

An example of a pair of concepts from the high-similarity group is shown in Figure 6-2. The F, B and S codes for each concept are shown in Table 6-2. In the rightmost column, a value of 1 refers to a match and 0 refers to a mismatch. The Jaccard coefficient is the number of codes shared by both sets (n=3) divided by the number of codes in either set (n=5) = 0.6 (the greatest Jaccard coefficient for pairs in this design task was 0.67). Figure 6-3 visualises the matching and mismatching codes associated with each concept.

The output of the pair creation process was 16 pairs of concepts, including one high-similarity and one low-similarity pair from each of the 8 design tasks.

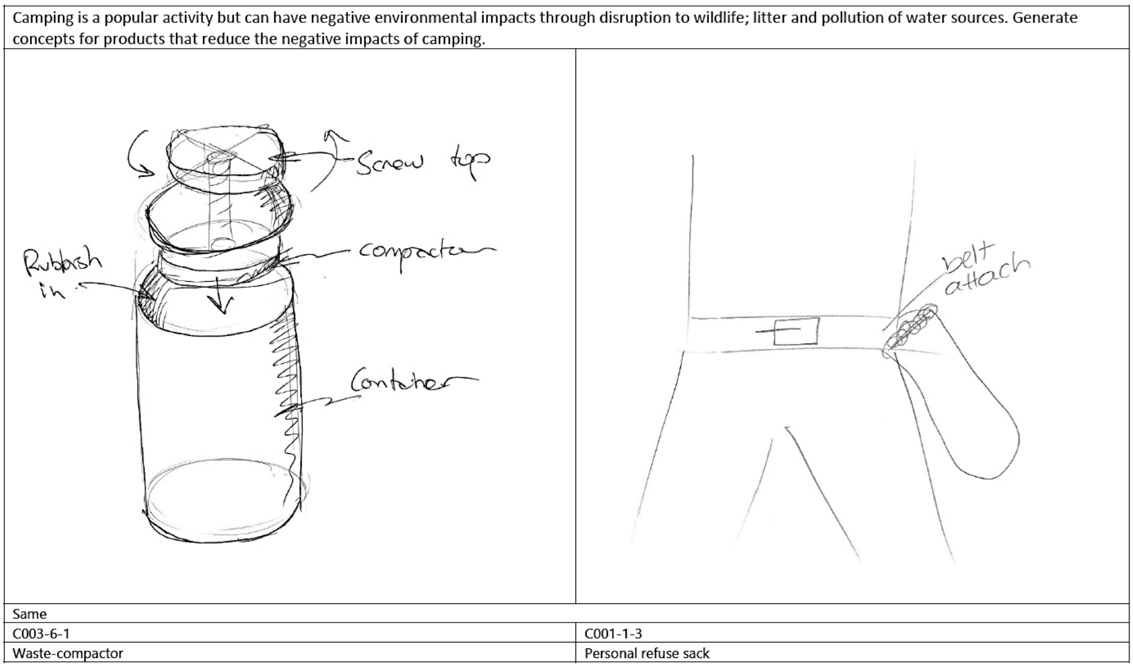


Figure 6-2 – An example of a pair of design concepts in the high alignability group.

Table 6-2 – The codes applied to the two design concepts in Figure 6-2 and the resulting Jaccard coefficient.

Code	Concept A (C003-6-1)	Concept B (C001-1-3)	Jacquard
F	F06-05 Facilitate storage of waste during camping trips	F06-05 Facilitate storage of waste during camping trips	1
B	B06-01 Compact rubbish	-	0
	B06-02 Contains rubbish	B06-02 Contains rubbish	1
S	S06-13 Rubbish receptacle	S06-13 Rubbish receptacle	1
	S06-10 Mechanical compacting device	-	0
		Jaccard Coefficient =	0.60

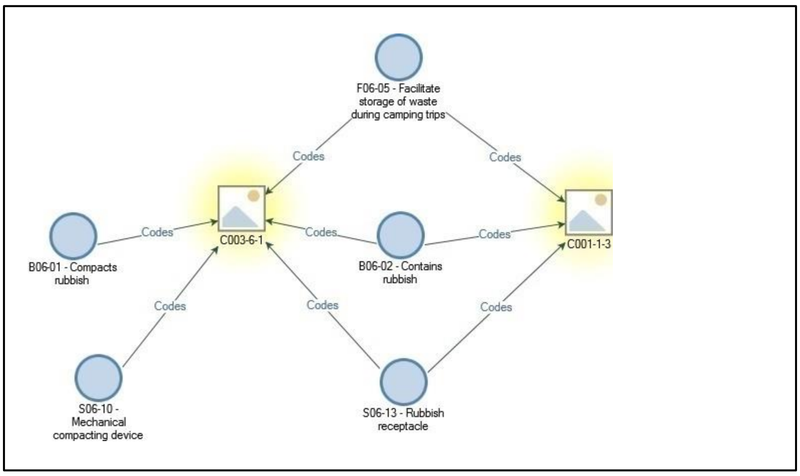


Figure 6-3 – A visualisation of the common and different codes applied to the design concepts in Figure 6-2.

6.2.3 Evaluation

To evaluate Stimuli Set A, the similarity ratings from the six participants in Sim-P1 were analysed to determine:

(i) whether the mean similarity ratings for the high-similarity group were significantly higher than for the low-similarity group, and

(ii) whether there were significant differences in the similarity ratings for the pairs *within* the high and low similarity groups.

The mean similarity ratings are shown in Table 6-3. Note that the Likert-type scale used in this experiment ranged from 1 (high similarity) to 7 (low similarity). The ordering of this scale confused participants and so the qualitative labels were reversed in subsequent experiments (i.e., 1 was low similarity). To maintain consistency with the rest of the experiments reported in this chapter the values have been inverted and thus in Table 6-3 higher numerical values represent greater similarity.

Table 6-3 – Mean (M) and Standard Deviation (SD) for similarity ratings

	High similarity			Low similarity			
Design task	**J**	**Mdn**	**n**	**J**	**Mdn**	**n**	**M - M**
DT03	0.67	5	6	0.10	3	6	2
DT06	0.67	2.5	6	0.10	1.5	6	1
DT07	0.67	6	6	0.11	3	6	3
DT08	0.60	5.5	6	0.14	3	6	2.5
DT09	0.75	3.5	6	0.10	5	6	-1.5
DT12	0.80	3.5	6	0.08	3.5	6	0
DT15	0.80	6	6	0.08	2	6	4
DT20	0.86	5.5	6	0.07	4.5	6	1
All		5.5	48		3	6	

J = Jaccard Coefficient, M = mean, SD = standard deviation, M-M = The difference between M for high similarity and low similarity.

A Mann-Whitney U test was run to determine if there were differences in the rated similarity of the base concepts in the high and low similarity conditions. Rated similarity was statistically significantly higher in the ‘high similarity’ group (*Mdn* = 5.5) than the ‘low similarity’ group (*Mdn* = 3.0), U = 521, z = -4.686, p < .001. The rightmost column of Table 6-3 shows the differences between the mean similarity ratings for the high- and low-similarity groups. This shows that even though *overall* the ratings of the high-similarity group were significantly higher than those of the low-similarity group, in one design task the opposite was true (DT09) and in another, there was no difference (DT12).

Kruskal-Wallis tests were conducted to determine whether there were differences in the similarity ratings for the stimuli in each group. For the 'high similarity' stimuli, median similarity ratings were statistically significantly different for different design tasks, $\chi^2(7) = 20.610$, $p = .004$. The same was true for the 'low similarity' stimuli, $\chi^2(7) = 15.319$, $p = .032$. No post-hoc tests were conducted for either condition.

From these results, it is inferred that the 'high similarity' pairs of design concepts were, as intended, rated as being more similar than the 'low similarity' pairs. However, the significant differences within each group indicate that the stimuli within each group were not homogeneous.

6.2.4 Discussion

It was proposed that if manipulation of the number of common F, B and S variables could be used to create pairs of design concepts belonging to internally homogeneous but externally distinct levels of similarity, then it could be considered a viable method for creating stimuli for use in the empirical research. The results show that the method provides a coarse-grain level of control over the mean value of rated similarity but that the high and low similarity groups are not internally consistent nor consistently different from each other. Thus, it did not pass the evaluation criteria.

The main implication of failing to meet the evaluation criteria was that the stimuli used in Sim-P1 were not suitable for use in evaluating the hypothetical model of design concept similarity judgements. This contributed to the decision to stop recruitment in Sim-P1 and move to Phase 2 of the research.

Although there was scope to attempt to improve the method, two issues highlighted the need for fundamental revisions to (rather than incremental improvement of) the method. One issue was that the FBS coding scheme used in Stimuli Set A did not differentiate between the problems that the concepts addressed, thereby failing to capture the problem-focused nature of the design brief that the participants responded to. The second issue was that coding behaviour as a standalone variable was difficult owing to the abstract and undetailed nature of the design concepts.

6.3 Phase 2: stimuli set B

Stimuli set B was created via a modified version of the concept-categorisation method described in Section 6.1.4. Design concepts were organised into concept ontologies and the distance between those design concepts was taken as a proxy for their similarity. The method was evaluated using the similarity judgements from Sim-P2. The stimuli were created to satisfy the minimum requirement of pairs of concepts that span a range of rated similarity (Section 4.2.4). An additional goal for the method was the creation of five levels of similarity. Although the evaluation criteria for this goal were not met, the stimuli did meet the minimum requirements.

6.3.1 Theoretical basis

Stimuli set B was created based on the theoretical assumptions held during phase 2 of the research. As per phase one, it was assumed that featural similarity was the predominant form of similarity and that manipulation of the common and different features of pairs of design concepts could grant *a priori* control over human similarity judgements. Assumptions about design concept representation were changed following the evaluation of Stimuli Set A. Specifically, there was a need to distinguish between the purpose and function of a design concept. This modification was also influenced by developments in the method for analysing problem-focused creative design ideation (Hay *et al.*, 2020) and is thus partly attributable to efforts by Dr. Hay.

The method used to create stimuli set B was a modified version of the concept categorisation method (Section 6.1.4). The intention was to organise design concepts in a concept ontology and use category location as a proxy for similarity. In line with a prior implementation of this method (Markman and Gentner, 1993b), it was assumed that design concepts that are close in a concept ontology are "more easily alignable" (Markman and Gentner, 1993, p.523) than those that are distant in the ontology, and would thus be rated as relatively more similar by designers.

Initial attempts were made to sort sets of design concepts into taxonomic hierarchies, but this proved not to be possible. The first approach was to use the same kind of taxonomy as shown in Figure 6-1 (Markman and Gentner, 1993b). This was not possible because design concepts are all members of the tangible, non-living things class, making the other categories in the classification irrelevant. The second approach was to use a smaller taxonomy with e.g., 3 hierarchical levels, and to further decompose tangible non-living things, hereafter 'artefacts', into additional subordinate categories. An example of such a hierarchy is shown in Figure 6-4.

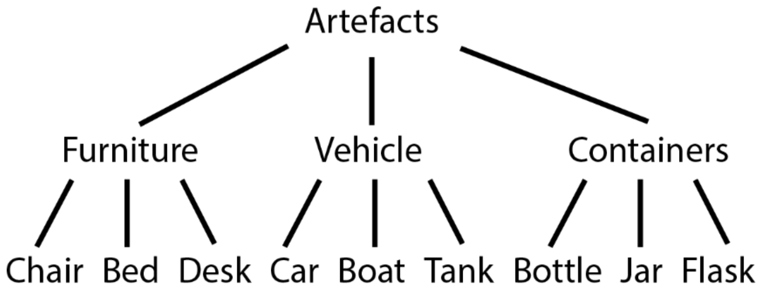
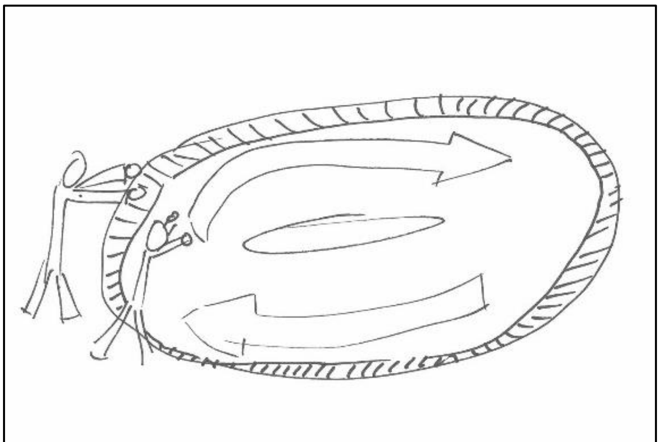


Figure 6-4 – An example of a three level taxonomic hierarchy of artefacts.

The problem with attempting to organise design concepts into taxonomic categories, as above, is that their novelty means that they cross category boundaries. To illustrate, consider the example in Figure 6-5. A design student generated a concept for a method of transporting water over long distances. The figure shows a "large water balloon that is a hoop so when pushed, it travels like a tank track". It could be argued that, if this were a product, it could be either a vehicle or a container. Decisions about which superordinate category it belongs to could influence the extent to which relative position in a hierarchy is useful for predicting perceived similarity.



Caption provided by designer: "A large water balloon that is a hoop so when pushed, it travels like a tank track. Thin tank tracks are on the outside to protect innards from rough terrain."

Figure 6-5 – An example of a design concept created in a concept generation experiment (Appendix 3B) and used in stimulus Sb_27.

This issue has been raised by Taura (2016, p.61), who theorised that design concepts cross category boundaries and thus do not share the same kind of conceptual structure as taxonomies of object concepts (Figure 6-6).

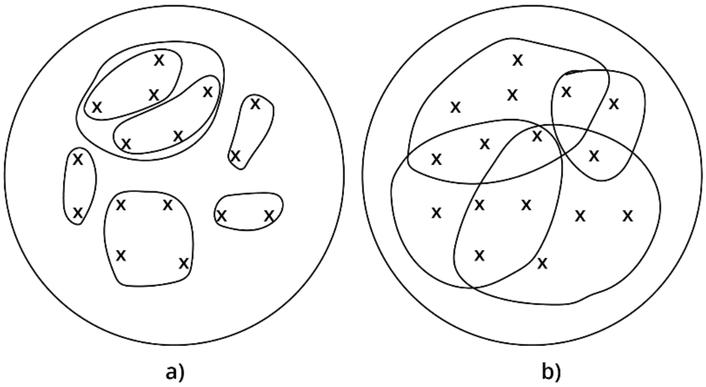


Figure 6-6 – A comparison of the conceptual structure of (a) existing products in taxonomic categories, and (b) design concepts in multiple taxonomic categories, redrawn from Taura (2016, p.61).

In light of the apparent challenges with organising design concepts into *taxonomic* classifications, an alternative approach was used to attempt to create classifications that were more suitable for design concepts. It has been proposed that design concepts can be classified in terms of common structure, behaviour, function or purpose (Rosenman and Gero, 1998). Although it was not clear whether design concepts can be categorised in a structure that reflects the relative similarity of the category members, a method was trialled to see whether the position of design concepts in such a classification would reflect perceived similarity.

A new scheme for design concept representation was defined based in part on the systematic approach for analysing exploratory design ideation (Hay *et al.*, 2020). It was proposed that it may be possible to categorise design concepts created in response to problem-focused briefs according to their purpose, function and need. The variables used to describe design concepts are defined in Table 6-4[15].

- **Problem and solution.** The scheme makes a distinction between a problem space and a solution space as per (Hay *et al.*, 2020). In response to problem-focused design briefs, designers identify and address a range of different sub-problems.

[15] These variables were originally given different names, as published in McTeague *et al.*, (2018). The variable names have been changed to more accurately reflect the intentions of the researcher at the time of creation.

Each designer will create a solution to one decomposed problem. The *purpose* of the design artefact must be to address the *problem* specified in the brief (or else the task has not been completed successfully).

- **Problem.** Designers decompose the problem, and the concepts created by multiple designers can be classified according to points in the problem space. This is assumed to occur through problem decomposition in which designers identify possible causal factors for the initial problem stated in the brief. For example, given the problem of 'unnecessary food waste', the designer may identify the causal issue of *"food becomes waste because it expires before it is used", or* further still *"food expires before it is used because the user is not aware that it is expiring".* Because of 1, decomposed problems have corresponding decomposed purposes. Each purpose is enabled by a function with a means of fulfilling that function.
- **Solution variables.** The solution variables describe the artefact and its purpose (utility in the context of the brief). The purpose of a design artefact in the context of a problem-focused design brief is to address a problem (or sub-problem). The function of the artefact enables the purpose. The function is achieved by the 'means' of fulfilling that function. Although design artefacts will typically have multiple functions and sub-functions, each addressed by a different means, it is assumed for categorisation that design concepts can be grouped and distinguished at the function level by their 'main function' and distinguished at the means level similarly.

Figure 6-7 shows an illustrative example of design concepts organised using this scheme. The numbers denote the five levels of ontological distance. The variables in the tree correspond to the 'solution variable' column of Table 6-4. $P_{(B)}$ = Purpose(brief), $P_{(-1)}$ = purpose (decomposed 1 level), $P_{(-2)}$ = purpose (decomposed 2 levels), F = function, M = means.

Table 6-4 – Variables used to describe the design problem and solutions in Stimuli Set B.

Variable	Description	Variable	Description
Problem (brief)	The problem stated in the design brief, e.g., 'unnecessary food waste'.	**Purpose (brief)**	The purpose of an artefact is its intended utility in addressing the problem, with respect to human utility. Given the problem (left), the purpose of the artefact is specified in the brief, i.e., to 'reduce unnecessary food waste'.
Problem (-1)	The maximum level of problem description that is not already contained within the design brief. For example, in the context of unnecessary food wastage, the problem may be: *"food becomes waste because it expires before it is used".*	**Purpose (-1)**	The purpose of the artefact decomposes with the problem. Given the problem (left), then the purpose of the artefact may be to: *"reduce the likelihood that food expires before it is used".*
Problem (-2)	A sub-division of the problem. For example, with the problem of food expiry above, the sub-problem may be: *"food expires before it is used because the user is not aware that it is expiring".*	**Purpose (-2)**	The sub-purpose is to address the sub-problem. Given the sub-problem on the last, the purpose of the artefact may be to: *"make the user aware of the expiration state of the food".*
		Function	The function describes what the artefact does in absence of the utility to the human being. For example, a refrigerator may display information about the upcoming expiry date of food.
		Means	A general term for the solutions to a function (Andreasen *et al.*, 2015, p.284). For example, the means of enabling the function (above) would be a refrigerator with some kind of date or food condition sensor and a display screen.

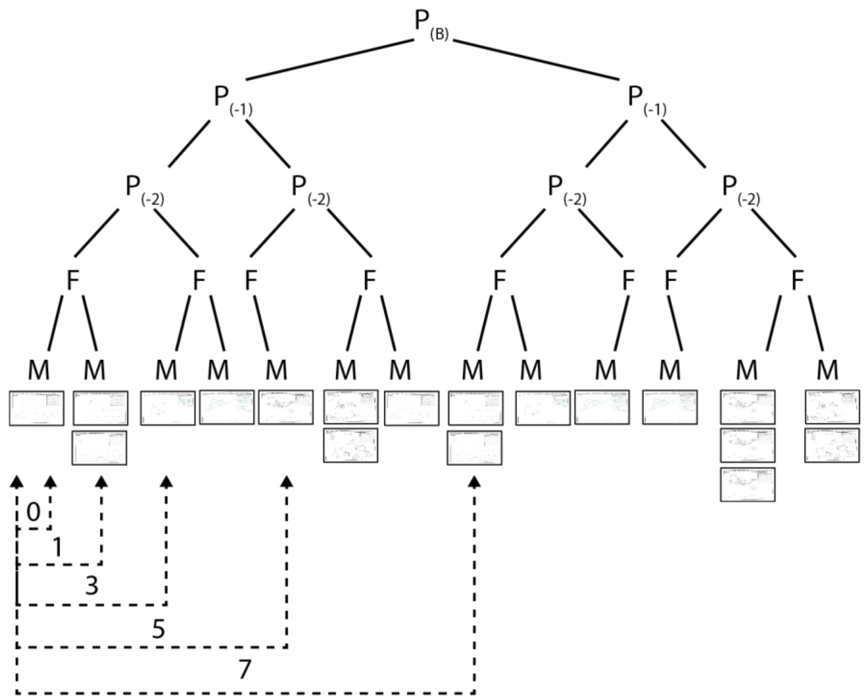


Figure 6-7 – An illustration of design concepts organised into hierarchical concept ontologies.

6.3.2 Method

Design concepts from four of the design tasks in the behavioural study sample were filtered, organised into categories, and pairs were created at the five levels of ontological distance.

Sample. Four design tasks were selected by the researcher from the ten used in the behavioural study (Table A3-5), excluding DT14 and DT19 which were deemed beyond the scope of the PDE domain in Phase 1 (Section 6.2.2). The behavioural study dataset was cloned at the same timepoint as in stimuli set A, and thus contained design concepts generated by 36 participants. The dataset is summarised in Table 6-5.

Table 6-5 – The four design tasks taken from the behavioural study (Appendix 3B).

Task #	Description	Tot	Exc	Used	Amb
DT03	Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.	86	9	61	16
DT06	Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.	75	20	47	8
DT07	Long-distance water transportation may be necessary in drought-prone developing nations but can be problematic due to a lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.	73	34	38	1
DT09	Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.	85	12	56	7

Tot = total number of design concepts associated with that design task, Exc = excluded, Amb = ambiguous, Used = design concepts used in the creation of Stimuli Set B.

Exclusion criteria. A new set of exclusion criteria were defined, removing the subjective criterion for sketch quality and more clearly defining the rules or exclusion. The number of excluded concepts is shown in the 'excluded' column (Exc) (Table 6-5). Concepts were excluded if they met any of the following criteria.

1. The concept is not a physical artefact, i.e. they were wholly an infrastructure or service solution.
2. The concept does not address the brief.
3. The designer identified a serious flaw in their concept and annotated this flaw in their response.
4. The concept did not have both a sketch and accompanying annotation,
5. The design concept could not be understood upon the first inspection. For example, when viewing the sketch and reading the annotations and accompanying descriptions, if it was not clear what the original creator was trying to communicate then the concept would be excluded.
6. The means or function of the product could not be interpreted.

Categorisation, filtering and exclusion. The researcher categorised the design concepts based on the definitions of purpose, function and means provided in Table 6-4. Design concepts were grouped in a top-down fashion, grouping concepts with similar purposes and then splitting the group according to sub-purpose, function and finally means. This resulted in a maximum of four hierarchical layers, however, in some instances, there was no need to divide a purpose into a sub-purpose. Concepts were classified as 'ambiguous' if there were multiple categories into which they could be placed at any level. The number of ambiguous concepts is shown in the 'ambiguous' column (Amb, Table 6-5).

Pair creation. Pairs of design concept sketches were created by selecting concepts according to their ontological distance in the hierarchy (Table 3). No design concept was used in more than one pair. A total of 40 pairs of concepts were created: 2(pairs) x 5(levels of ontological distance) x 4(design tasks). Examples are provided of pairs of design concepts at ontological distances 1, 3 and 5 in Appendix 4A.

6.3.3 Evaluation

Stimuli set B was evaluated against two criteria using the similarity ratings from Sim-P2. 11 Participants provided similarity ratings on a 1-9 scale with 1 being the least similar and 9 being the most similar.

- The first criterion was whether the stimuli spanned a range of similarity ratings. Satisfaction of this first requirement was necessary if the data from Sim-P2 were to be used to address the research questions associated with the hypothetical cognitive model of design concept similarity judgements.
- The second criterion was whether the stimuli in the five levels of ontological distance were internally homogeneous and externally distinct. Satisfaction of this criteria would mean that there were no statistically significant differences between stimuli at the same level, but that there were significant differences between different levels.

The similarity explanations provided by participants in Sim-P2 were examined to evaluate the concept representation scheme. The researcher analysed the explanations to search for examples that would contradict the assumptions of the design concept representation scheme (P, F and M variables).

6.3.3.1 Criteria 1: pairs of concepts spanning a range of similarity ratings

The minimum requirement of the stimuli that would facilitate their use in Sim-P2 was that the pairs of design concepts span a range of similarity ratings. Figure 6-8 shows the boxplots for rated similarity for all pairs and the 10 pairs from each design task, arranged in descending order. Visual inspection of the boxplots demonstrates that the stimuli span a range of median similarity ratings from 1-9 for the whole sample of concepts (Figure 6-8, a). A similar pattern is found for the pairs in each design task other than those from DT09 (Figure 6-8, b-e). This demonstrated the satisfaction of the minimum requirements for stimuli creation and the utility of Stimuli Set B for use in Sim-P2. The distribution of the median similarity ratings for each pair across low (1-3.33), medium (3.34 – 6.66) and high (6.67 – 9.00) similarity by median value is n = 14/9/17. This reflects an approximately symmetric, bimodal distribution for all of the similarity ratings (Figure 6-9).

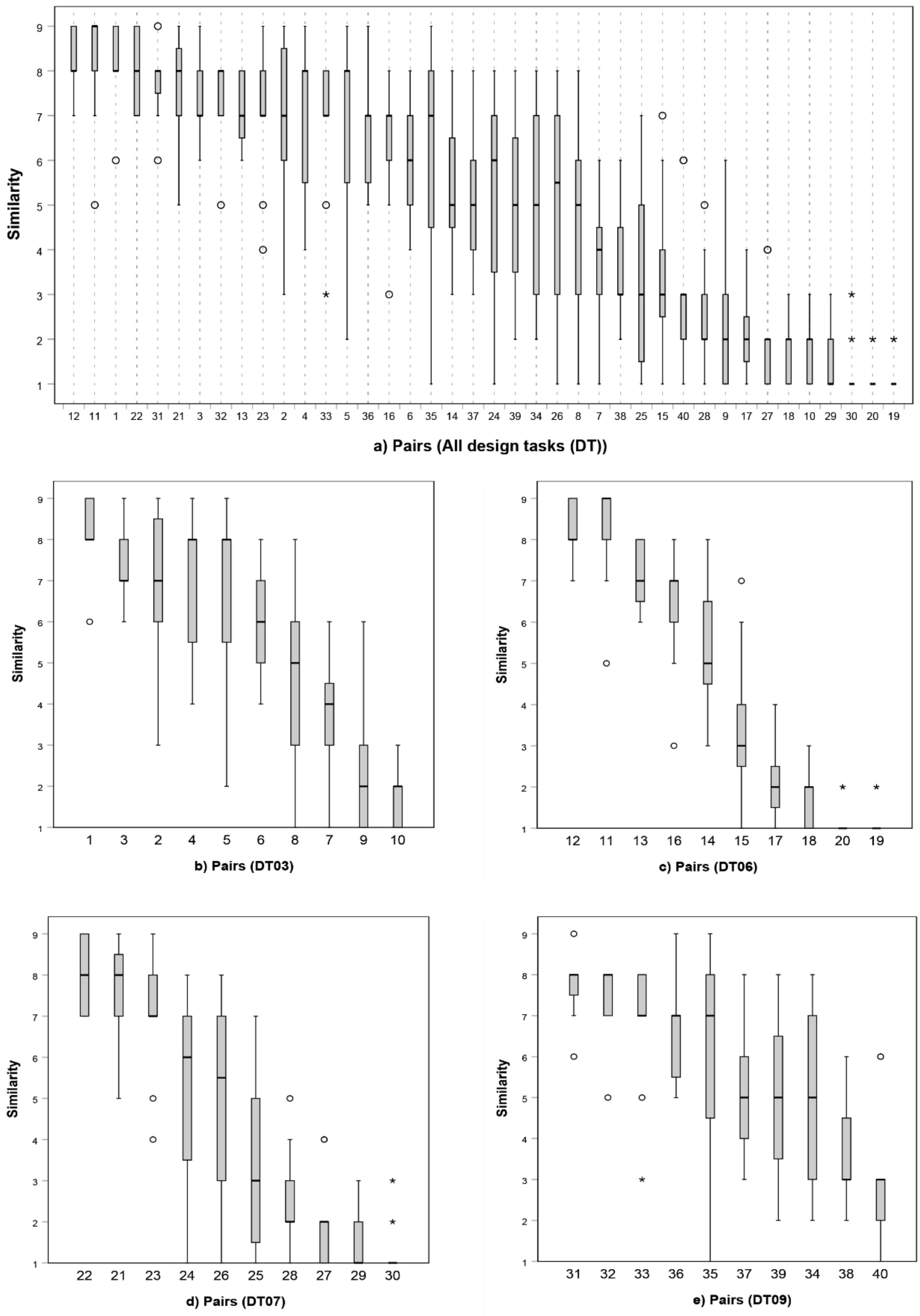


Figure 6-8 - Boxplots of similarity ratings of Stimuli Set B in descending order for a) all design tasks, and (b-e) individual design tasks.

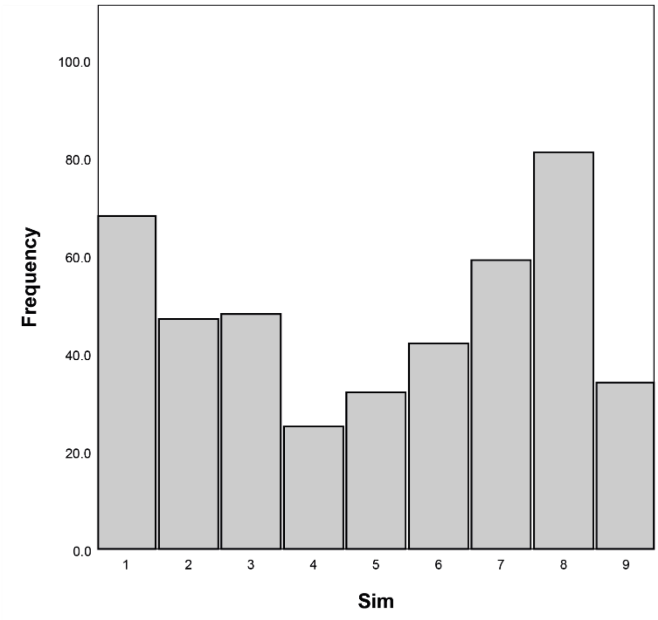


Figure 6-9 – Histogram showing frequencies of similarity ratings from Sim-P2.

6.3.3.2 Criteria 2: ontological distance as a determinant of similarity

The first set of criteria for stimuli set b concerned the relationship between ontological distance and similarity. This was to test whether the similarity ratings for the stimuli on each of the five levels of ontological distance were (i) homogeneous at the same distance and (ii) significantly different between levels. Similarity ratings were taken from the similarity rating task in Sim-P2.

Table 6-6 – Summary data for Stimuli Set B. Median and mean rank similarity for 8 stimuli across 4 design tasks.

		DT03		DT06		DT07		DT09		All
OD		A	B	A	B	A	B	A	B	
	Mdn	8.0	7.0	9.0	8.0	8.0	8.0	8.0	8.0	8
0	MR	52.64	34.14	55.41	56.45	42.27	44.05	39.50	31.55	349.93
	n	11	11	11	11	11	11	11	11	88
	Mdn	7.0	8.0	7.0	5.0	7.0	6.0	7.0	5.0	7.0
1	MR	57.45	53.64	53.14	30.68	53.68	29.05	49.45	28.91	277.48
	n	11	11	11	11	11	11	11	11	88
	Mdn	8.0	6.0	3.0	7.0	3.0	5.5	7.0	7.0	6.0
3	MR	57.09	50.09	22.55	53.00	23.00	39.50	52.27	54.09	233.21
	n	11	11	11	11	11	10	11	11	87
	Mdn	4.0	5.0	2.0	2.0	2.0	2.0	5.0	3.0	3.0
5	MR	53.00	59.18	29.18	24.75	24.09	34.30	68.50	50.86	138.32
	n	11	11	11	10	11	10	10	11	85
	Mdn	2.0	2.0	1.0	1.0	1.0	1.0	5.0	3.0	1.0
7	MR	49.09	43.00	28.73	28.73	37.27	30.14	76.59	62.45	90.98
	n	11	11	11	11	11	11	11	11	88

DT = design task, OD = ontological distance, Mdn = median, n = number of similarity judgements, MR = mean rank. MR is calculated within each level except in the 'all' column where it is calculated across levels.

Similarity and ontological distance. A Kruskal-Wallis test was conducted to determine if there were differences in the rated similarity between design concepts at different ontological levels (mean ranks data in Table 6-6, 'All' column). Distributions of similarity scores were not similar across the five groups, as assessed by visual inspection of the boxplot (Figure 6-10). The mean rank of the similarity ratings differed from distance 7 (mean rank = 90.98), to 5 (mean rank = 138.32) to 3 (mean rank = 233.21) to 1 (mean rank = 277.49) to 0 (mean rank = 349.93). Similarity scores were statistically significantly different between the different levels of distance, $\chi^2(4) = 244.96$, $p < .001$. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons (adjusted p-values are presented). There were statistically significant differences in similarity ratings between all pairwise comparisons ($p \leq .001$) other than for distances 7-5 ($p = .127$) and 3-1 ($p = .190$).

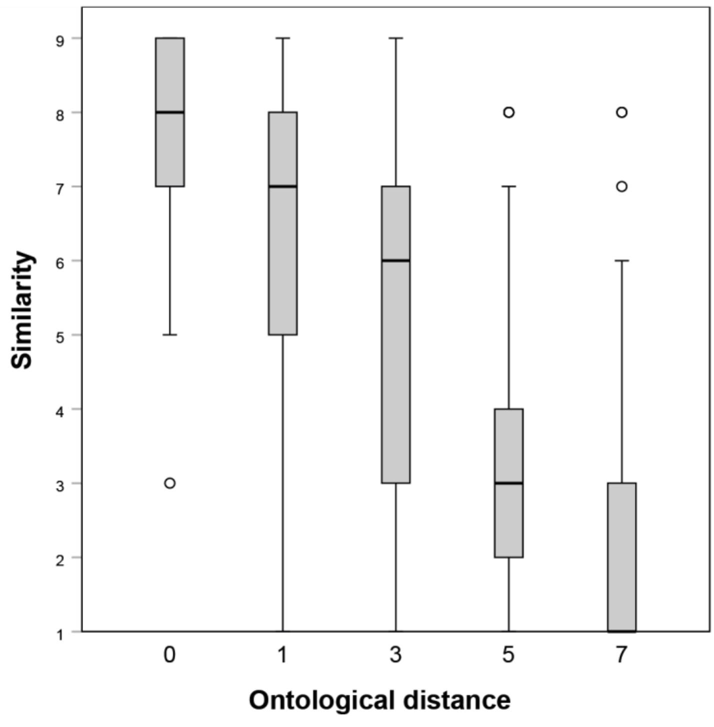


Figure 6-10 – Boxplot of similarity ratings at five levels of ontological distance.

Within-level homogeneity. A Kruskal-Wallis test was conducted to determine if there were differences in the rated similarity between the pairs of design concepts at each level of ontological distance (mean ranks data in Table 8-6). Table 6-6 shows the mean similarity ratings for each of the 40 pairs of design concepts. The letters A and B are arbitrary designations that denote the two pairs from each design task at the same ontological distance. Five tests were conducted, one for each level. Distributions of similarity scores were assumed not to be similar across any of the five distance levels, based on visual inspection of the boxplots (Figure 6-11). The similarity ratings were statistically significantly different between the design concepts within each level for distances 1 ($\chi^2(7) = 19.609$, $p = .006$), 3 ($\chi^2(7) = 24.287$, $p = .001$), 5 ($\chi^2(4) = 36.177$, $p < .001$) and 7 ($\chi^2(4) = 42.049$, $p < .001$), but not for distance 0 ($\chi^2(7) = 11.829$, $p = .106$).

The results of the two sets of statistical tests show that stimuli set B fails the two evaluation criteria.

(i) **Similarity at different ontological levels.** The similarity ratings for the pairs of design concepts are not significantly different for different levels of ontological distance. Specifically, the numerical difference observed between levels 7 and 5, as well as 3 and 1, could be attributed to chance.

(ii) **Similarity at the same ontological level.** There were significant differences between the similarity ratings for pairs of design concepts within the same ontological level for four out of the five levels. This demonstrates poor homogeneity for concepts that were intended to be equivalently similar.

These findings led to the decision to conduct a second evaluation based on the minimum requirement that the sample of design concepts span a range of rated similarity.

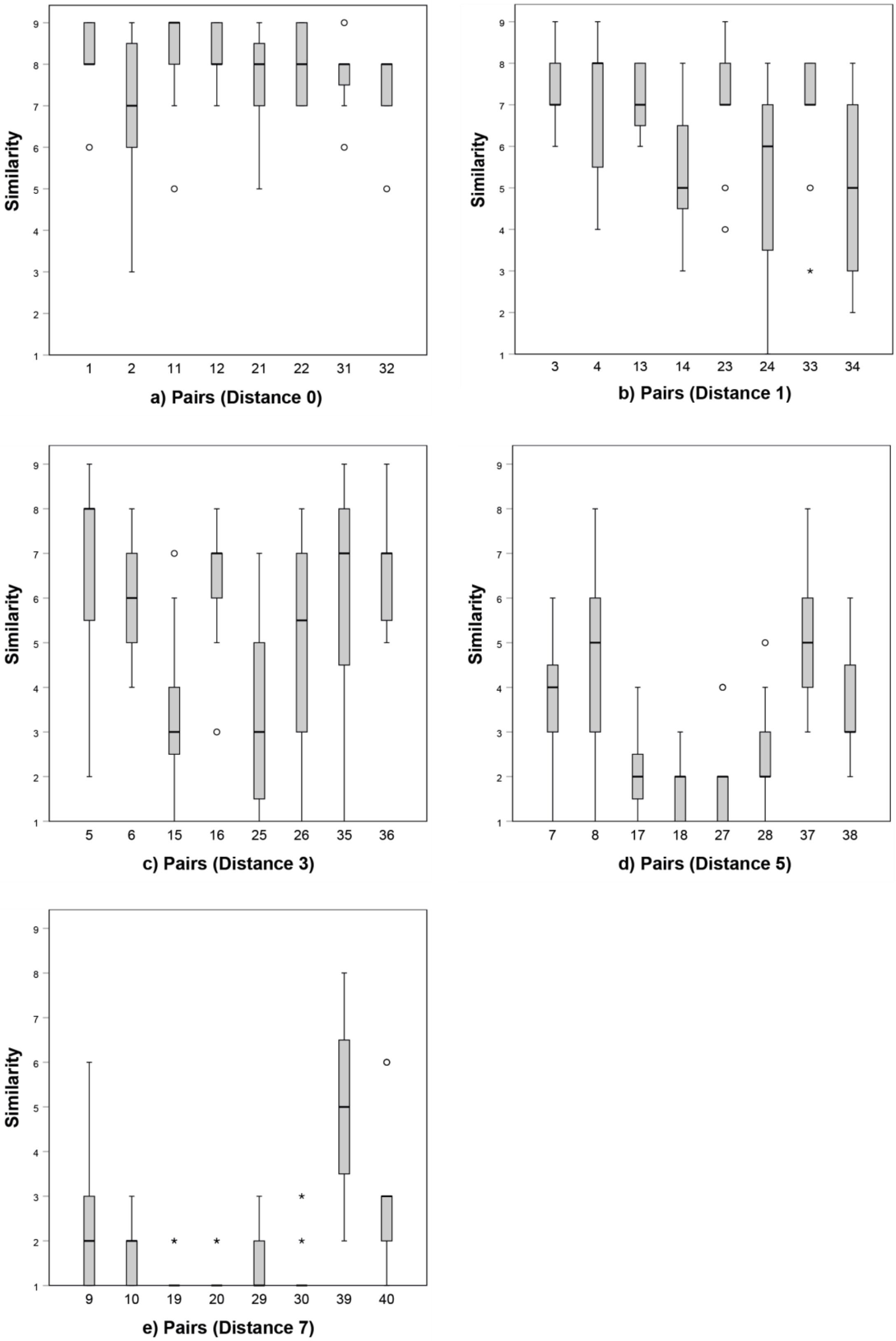


Figure 6-11 - Boxplots of similarity ratings for each stimulus (pair of design concepts) at five levels of ontological distance.

6.3.3.3 Insights from similarity explanations

The final evaluation of the stimuli creation method was conducted by examining the explanations for similarity ratings provided in Sim-P2. The purpose of the analysis was to identify the rationale for similarity judgements that contradicted the assumptions of the design concept representations scheme used in Phase 2. Five issues are presented here that are inconsistent with the PFM design concept representation scheme. They are inferences made by the researcher based on the explanations (Table 6-7). The analysis was not exhaustive nor was it conducted to identify the frequency of the counterexamples, only that they existed. Preliminary analysis of these findings has been published previously (McTeague *et al.*, 2018) but some inferences have been updated since then.

Table 6-7 - Inferences drawn from the similarity explanations

#	Concept (left)
1	**Participants are aware of abstract properties of design concepts such as morality and complexity.**
A	"both make seating uncomfortable for user. one less sadistic"
B	"The pull is a simple pull string, the push is a more complex handle that attaches to the barrel."
2	**Participants are sensitive to the taxonomic class of category concept into which a design concept may belong.**
A	"Both are tents but waste-reducing aspects of the two concepts are entirely different"
3	**Similarity judgements may be focused on the user rather than the artefact.**
A	"Both require resistance work; however, one targets a specific muscle group, whilst the other attempts to offer a more general solution which enables working various body parts."
4	Participants identified elements of service systems in the design concepts
A	"Concept 1 a service + application based product; Concept 2 is a physical product."
5	Participants commented on concepts having different numbers of functions
A	"Both have exact same function - tracks and displays food date, but one has additional function of keeping food fresh for longer"

Inferences made by the researcher are numbered (1-5). Examples that support each inference are designated with letters.

These five inferences have implications for the design concept representation scheme and the filtering applied in the process of selecting design concepts.

- Inferences 1 and 2 suggest that participants use variables other than the purpose, function and means variables in their similarity judgements. Example 1A shows that a participant has made a similarity judgement based on an affective judgement about the degree of 'sadism' inferred to exist in the product. Such a dimension is not explicitly accommodated by the P, F or M variables although may be an abstraction of the function or purpose of the artefact (the product in question provides negative feedback for not completing exercises).

- Inference 3 raises the question of the extent to which an artefact-centric model of concept representation is sufficient for capturing salient features relating to the user, rather than the artefact. Example 3A is an explanation for concepts that can facilitate exercise in seated positions at home or in the office. The participant's similarity judgement focuses on the muscles that are activated when using the product. This may be captured in the purpose of the artefact, i.e., that it 'allows a user to exercise a specific muscle group when seated'. However, it is unclear whether the PFM scheme is sufficient for capturing similarity derived from user-centred commonalities and differences.
- Inference 4 highlights issues with the design concept filtering process. Although concepts were excluded if they were 'wholly' service systems solutions, it appears that some concepts remain that *contained* service systems and that participants are sensitive to this in their similarity judgements.
- Inference 5 shows that participants are sensitive to the *number* of e.g., functions, rather than what those functions are *per se.* Jameson *et al.* (2005) propose that *number* is treated as an alignable difference in comparison.

These five inferences are discussed in the context of the two prior evaluations in the subsequent discussion.

6.3.4 Discussion

Stimuli set B satisfied the requirement for pairs of design concepts that span a range of similarity ratings (Section 6.3.3.1) but failed to meet the criteria set for multiple levels of ontological distance (Section 6.3.3.2). Because the stimuli satisfied the minimum requirements for use in the similarity model evaluation, the similarity judgements elicited in Sim-P2 were used to evaluate the research questions associated with the hypothesised Dual-Process model of design concept similarity judgements.

Despite the failure to meet the evaluation criteria for the five levels of ontological distance, the method demonstrated a degree of *a priori* control of similarity ratings. There were significant differences between four of the five levels of ontological distance, suggesting that category location may be associated with design concept similarity to some extent. In this respect, the method was more successful than prior implementations in the psychology literature. Markman and Gentner (1993) found that significantly higher similarity ratings were given to pairs of distance 0 than to pairs of distance greater than 0, but that there were no significant differences between any other levels. Notably, this study does not report the homogeneity of ratings from within the same level and so a more extensive comparison is not possible.

In light of these outcomes, it may be possible to improve the method to such an extent that it can be used to provide fine-grain control over the similarity of pairs of design concepts. Although this was not done as part of this thesis, future work may benefit from a discussion of methodological and theoretical issues that may determine the potential scope for improvement.

6.3.4.1 Methodological issues

Methodological issues may have limited or over-stated the success of the method. Two potential issues are the reliability of the categorisation method and the filtering process. The qualitative analysis in (Section 6.3.3.3) revealed at least one example of products that had service system elements in them. The filtering criteria for selecting design concepts were intended to eliminate concepts that were wholly service systems but did not capture all service system elements. This issue may have been avoided by conducting the filtering twice at separate time points or by eliciting the help of an independent judge in the filtering process.

This previous issue raises the issue of reliability more generally. As the entire concept creation process was carried out by the researcher there is a risk of bias in the categorisation process. For example, if the placement of design concepts into category locations was influenced by personal characteristics of the researcher, or by differences in e.g., mood and outlook over different days, then the positions of the design concepts in the classification (and thus the relative similarity of each pair) may incorporate these biases as extraneous variables. Steps taken to assess the reliability of the concept representation scheme would have necessitated more explicit operationalisation of the PFM scheme and may have improved the reliability and validity of the method.

6.3.4.2 Theoretical issues (concept representation)

Issues with the theoretical basis of the categorisation method may limit the scope of refinement that is possible with methodological improvements.

- The research in phase 2 was carried out with the assumption that featural processing is the predominant form of similarity processing. However, if some stimuli were predominantly (or additionally) processed thematically then this would explain the significant differences within each level of ontological distance.

- The PFM and variables do not capture all aspects of participant similarity ratings. This was to be expected given the relatively coarse granularity of the representation scheme; concepts at the same points in the ontology were assumed to share many common features but not to have identical features. The extent to which this leads to heterogeneity in participant responses is unclear. The similarity-explanations show that participants are sensitive to user-artefact interaction and abstract emotive or emergent properties. Although the five levels of ontological distance demonstrated some degree of *a priori* control over similarity ratings, these extraneous factors show that the scheme cannot capture all aspects of human similarity judgements. It is not clear to what extent this limits the utility of the concept categorisation method.
- Different design tasks may elicit similarity judgements based on different information within a design concept. The example of the similarity judgement that focuses on the user rather than the artefact (Table 6-7, 3), in conjunction with the inconsistent similarity ratings given for the exercise task (Figure 6-8, e), suggests that the design task itself may act as a confound for similarity judgements.

These issues were considered when establishing the theoretical basis for stimuli set C.

6.4 Phase 3: stimuli Set C

Stimuli Set C is a modified version of Stimuli Set B where more stringent inclusion criteria have been applied and 12 of the pairs have had at least one design concept swapped with equivalent concepts from the same ontologies used in Stimuli Set B. The previous set of 40 pairs from 4 design tasks has been split into a warmup set (10 pairs from 1 design task) and the main set (30 pairs from 3 design tasks). The pairs were shown to satisfy the minimum requirement of spanning a range of similarity ratings.

6.4.1 Theoretical basis

Stimuli set C was created by modifying stimuli set B. This was done because Set B had already been shown to span a range of rated similarity, but some limitations could be overcome by further filtering. Owing to the limitations of the purpose-function-means representation scheme, phase 3 was initiated without any explicit assumptions about the mental representation of design concepts. Rather, the Stimuli Set B was modified to address the limitations discussed in Section 6.3.4.

6.4.2 Method

Stimuli set C was created by making two kinds of modifications to set B. 10 pairs of design concepts were moved to create a warmup set and a new round of filtering and exclusion was conducted.

Creation of a warmup set. Stimuli set B comprised 40 pairs of design concepts (4 design tasks x 10 pairs). To create stimuli set C, all 10 pairs from DT09 – the seated exercise brief, were placed into a separate warmup booklet. DT09 was moved to a warmup because the distribution of similarity ratings (Figure 6-8e) was inconsistent with the pattern of descending similarity found in the pairs from the other three tasks. The remaining 30 pairs (3 DT x 10 pairs) remained in the 'main' set. This main set was then subject to the updated exclusion criteria. No replacements were made for the warmup set (DT09).

Exclusion criteria. The updated exclusion criteria are listed in Table 6-8, showing the criteria used for stimuli sets B and C. The original criteria were re-applied and three new criteria were added. The new criteria were added to further homogenise the design concepts and to bring them more in line with the kinds of products that the PDE participants would be familiar with.

- The criterion for excluding concepts that were 'wholly' service system or infrastructure changes was made more severe by including a second requirement that the concept does not *contain* any service system or infrastructure elements.
- Design concepts that had two sub-purposes were swapped out for equivalent concepts that had one sub-purpose.
- Concepts were excluded if the artefacts functioned by being acted upon by the environment, such as by 'being biodegradable'.

Table 6-8 – Filtering and exclusion criteria for Stimuli Set C

Description	**Stimuli set B**	**Stimuli set C**
The concept is not a physical artefact, i.e. they were wholly an infrastructure or service solution,	Y	Y
The concept contains elements of a service system or involves modifications to industrial supply changes to implement		Y
The concept does not address the brief,	Y	Y
The designer identified a serious flaw in their concept and annotated this flaw in their response,	Y	Y
The concept did not have both a sketch and accompanying annotation,	Y	Y
The sketch representation could not be deciphered, or	Y	Y
The means or function of the product could not be interpreted.	Y	Y
The design concept has two functions that address the brief and there is a similar single-function design concept in the sample.		Y
The artefact functions by being acted upon by the environment.		Y

Summary of changes. Table A4-6 contains a list of the stimuli in Stimuli Set C and shows the design concepts that were changed alongside the reasons for the changes. 14 design concepts from 12 pairs were swapped. The swaps were because the stimuli violated the requirements of no service system elements (5), no passive functions (4), no dual functions (3) and a lack of clarity (2). The main set of 30 pairs in stimuli set C has 18 pairs in common with set B.

6.4.3 Evaluation (similarity manipulation check)

The intended IV manipulation for Stimuli Set B was that the pairs of design concepts spanned a range of mean rated similarity across the sample of 30 pairs. Evaluation of the set was conducted by visual inspection of the similarity ratings taken from Sim-P3.

As intended, the stimuli spanned a range of similarity ratings along the 9-point scale. Figure 6-12 shows the boxplots of Sim(all) after outlier removal (see Section 0) for (a) all pairs, and (b) the ten pairs from each design task (b – d), arranged by descending median. This demonstrates that the stimuli set still adheres to the minimum requirement for stimuli creation after modification from stimuli set B. The distribution of the median similarity ratings for each pair across low (1-3.33), medium (3.34 – 6.66) and high (6.67 – 9.00) similarity by median value is n = 14, 6, 10. This reflects a bimodal distribution for Sim(all) with a higher frequency of low-similarity ratings across the set (Figure 6-13).

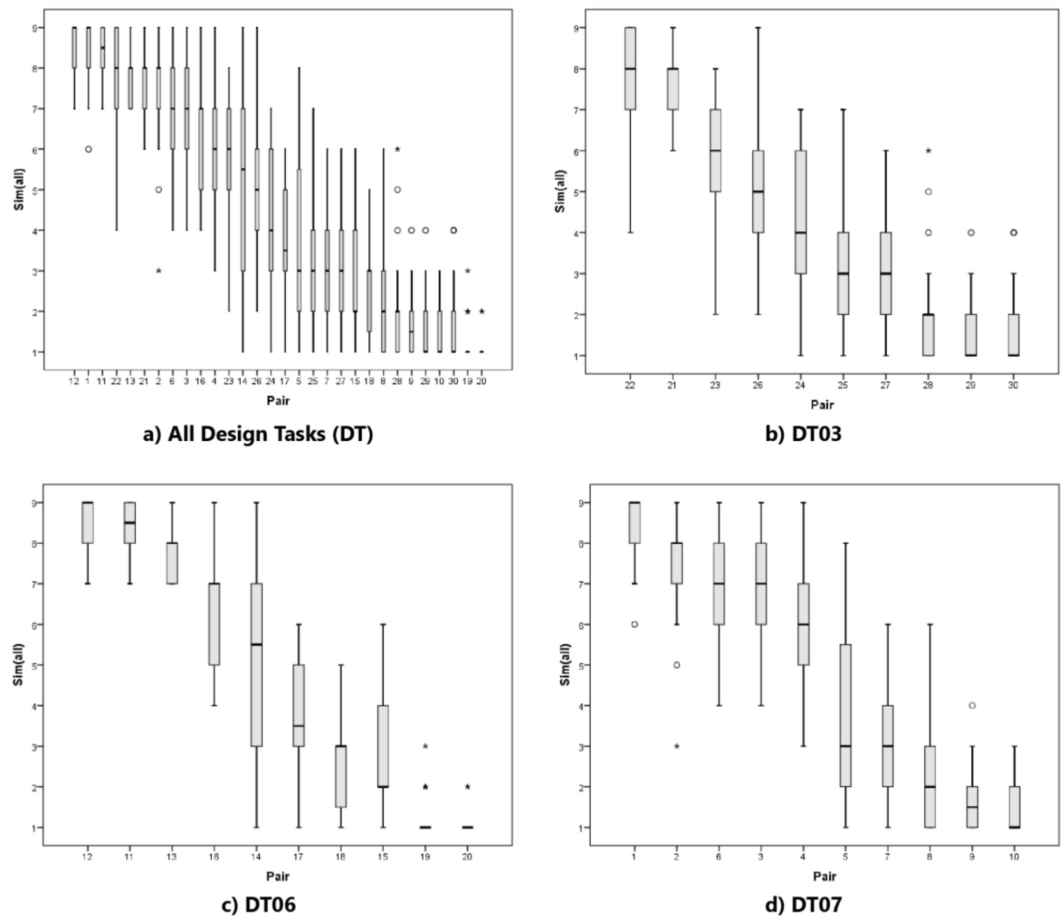


Figure 6-12 – Boxplots for Sim(all) ordered by descending median similarity, a) all design tasks, b) Design task 03, c) Design task 06 and d) Design task 07.

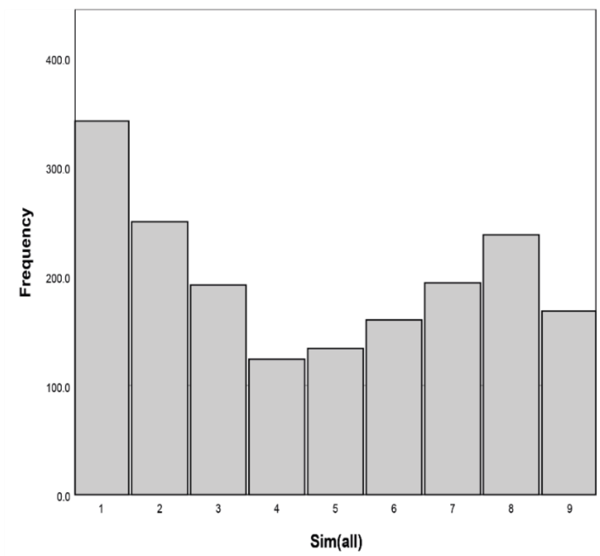


Figure 6-13 – Histogram showing frequencies of similarity ratings Sim(all) from Sim-P3.

6.4.4 Discussion

One implication of creating Stimuli Set C by modifying Set B is that the former will inherent any biases present in the latter. One possible risk is that Stimuli Set B contains a stimulus compatibility bias towards featural processing. Stimulus compatibility means that some stimuli may be more or less compatible with one kind of processing, i.e., featural or thematic. If Set B were more compatible with say, featural processing than thematic processing, then this may carry over to Set C and influence the results of the similarity and combination experiments in Phase 3. This was considered during the analysis and discussion of Sim-P3 and Combo-P3, but no problems were identified.

6.5 Chapter summary

This chapter has presented the development and evaluation of three stimuli sets A, B and C that correspond to the three phases of research 1, 2 and 3 as shown in the research map (Figure 4-6). The development of each set was presented, including the theoretical basis of the method, the procedure for creating the stimuli, their evaluation and a discussion of the strengths and weaknesses of each method.

- Stimuli Set A was created based on the assumption that the similarity of pairs of design could be controlled by manipulating the degree of common Function, Behaviour and Structure variables. The stimuli were created by coding the design concepts and similarity was computed by Jaccard's coefficient. The stimuli were used in Sim-P1 and Combo-P1. The stimuli set did not meet the success criteria and conceptual issues were identified in the stimuli creation method.
- Stimuli Set B was created based on the assumption that the similarity of pairs of design concepts could be manipulated based on the position of a design concept in a concept ontology. A different concept representation scheme was used this time, describing design concepts in terms of Purpose, Function and Means variables. The stimuli satisfied the minimum requirement of spanning a range of similarity ratings, based on the similarity ratings from Sim-P2.
- Stimuli Set C was created by modifying the stimuli in set B to preserve the range of similarity ratings and adhere to more stringent exclusion criteria. Again, the stimuli satisfied the minimum requirement of spanning a range of similarity ratings, based on the similarity ratings from Sim-P3. In a departure from the previous two phases, stimuli set C was made in advance of, rather than after the establishment of a design concept representation scheme.

The stimuli created through the methods in this chapter were summarised in the materials and methods chapter (Table 5-3)

7 RESULTS PT. I: DESIGN CONCEPT SIMILARITY JUDGEMENTS

This chapter presents the analysis, results and discussion of two experiments carried out to satisfy Obj. 3 - to *propose and test a cognitive model of design concept similarity judgements.* Testing of the model involved two steps, as shown in the research plan (Figure 4-5):

- Answer the research questions – *determine the number, type, and prevalence of cognitive processes involved in design concept similarity judgements.* This was carried out by the similarity experiment in phase 2 (Sim-P2).
- Test the hypotheses – *evaluate predictions of the Structural Alignment model of comparison-based similarity judgements.* This was carried out by the similarity experiment in phase 3 (Sim-P3).

Similarity Experiment 1 (Sim-P1) was conducted as a pilot study to test the experimental procedure for the similarity rating task and to facilitate the evaluation of Stimuli set A, so no results from this experiment are reported in this chapter. The results from Sim-P2 and Sim-P3 are discussed in Section 9.1 to evaluate the proposed Dual-Process model of design concept similarity judgements.

7.1 Similarity experiment 2 (Sim-P2): featural and thematic processing in explanations for similarity ratings

Similarity Experiment 2 (Sim-P2) was designed to test the research questions associated with the cognitive model of design concept similarity judgements. This involved answering three research questions (Table 7-1). The procedure for this experiment is reported in Section 5.2.2. Participants' written explanations for their numerical similarity ratings were analysed to determine whether their similarity judgements were 'feature-based', meaning that they were based on the common and different features of two concepts or 'thematic', meaning that they were based on the external and complementary relations between the two design concepts. Featural responses would indicate that similarity judgements occurred via a comparison process and thematic responses would indicate that similarity judgements occurred via a scenario creation process.

Table 7-1 – The research questions addressed in Sim-P2

Ref.	**Research question**
RQ – S1	Can the explanations for design concept similarity ratings be coded as feature-based and thematic?
RQ – S2	What is the relative prevalence of each explanation type?
RQ – S3	What is the relationship between concept pair similarity and the type of similarity explanation?

7.1.1 Results

7.1.1.1 Featural and thematic processing in similarity explanations

A total of 439 explanations were examined from 440 responses (one similarity rating was not accompanied by an explanation). Of 439 explanations, 412 were coded as 'feature-based', none were coded as 'thematic', and 27 were coded as 'other'. The researcher and an independent judge achieved 94.99% agreement on the entire sample (417/439) (Section 5.2.2.1). The 22 disagreements were resolved in arbitration. Of these, two were agreed to be featural responses, and the rest were coded as 'other'. An example of the similarity ratings and the explanations given for a pair of design concepts is shown in Figure 7-1 and Table 7-2 (p.177). All of these responses were coded as featural.

These results address both research questions. The agreement between two judges shows that explanations for similarity ratings can be coded reliably as featural explanations (RQ-S1). Since no thematic relations were found, the ability of judges to agree on thematic relations has not been demonstrated. All explanations for rated similarity that could be coded unambiguously were feature-based explanations (RQ-S2). As the responses were overwhelmingly featural, the third research question (RQ-S3) was rendered redundant.

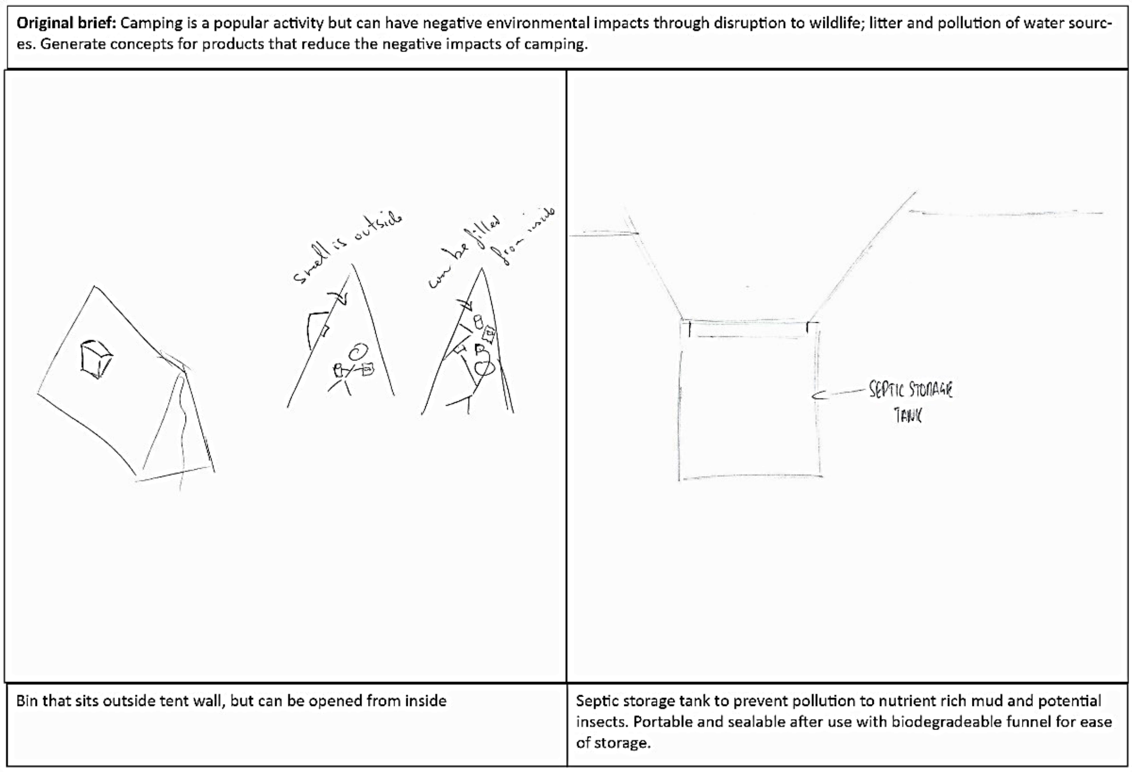


Figure 7-1 – Pair of design concepts Sb_15.

Table 7-2 – Examples of similarity ratings and explanations for stimulus Sb_15

ID	Commonalities	Sim
1	Concepts address different sub-issues.	1
2	Both contain waste (and smell) to prevent external pollution. Differences in the mechanism to open/close, containment and location of the device.	3
3	Both store waste	3
4	both concerned with waste produced by the campers	6
5	First one address a human issue the other an environmental issue. but they are both bins and for use in the wild so not too un similar.	4
6	Both seek to reduce ill effects of waste on environment; however, achieved by different mechanisms. one approaches rubbish, the other deals with something else.	2
7	both concern storage of waste	4
8	Similar that waste is enclosed, but how and the form of the product enclosing it is very different.	3
9	Both facilitate storage of rubbish during camping, though bin more simple, has less structural complexities	7
10	One looks at trash the other bodily waste	1
11	Both are bins but different approaches	3

ID = participant number, Sim = similarity rating (1-9 scale)

7.1.1.2 'Other' codes

The 27 responses coded as 'other' are listed in Table A7-2. None of these responses could clearly be labelled as feature-based or thematic. 23 responses are generic assertions or negations of high similarity or extreme difference. 4/27 responses do contain indications of thematic (external and complementary) relations but are sufficiently ambiguous that clear thematic processing cannot be inferred.

The 23/27 assertions of similarity or difference are listed in Table A7-2. Examples of these responses are: "different ideas", "no connection whatsoever" and "Very similar idea, but products delivering it are very different.". These responses do not refer to any specific features of the concepts nor do they make generic mentions about e.g. the function, behaviour or structure of the responses, and thus they cannot be coded as feature-based. Further, they do not include evidence of external or complementary relations between the two concepts and thus cannot be coded as thematic.

The four responses that could potentially be construed as thematic are listed in Table 7-3. The thematic aspects of the responses are highlighted in the subsequent list along with the rationale for not coding them as thematic.

- Response 24 – If both concepts were given the feature of 'being like an office chair', or if the participant has interpreted both concepts as being 'kinds of' office chairs, then this could be construed as a feature-based commonality. An alternative interpretation is that this response is neither featural nor thematic, but rather represents a unique kind of explanation in which two design concepts are placed into a scenario with a third entity, i.e., the category concept of an 'office chair'.
- Response 25 – The participant may have described an external and complementary relationship between the 'use-case' of one concept and the 'need' that the other concept addresses. However, the language is not sufficiently clear to unambiguously interpret the explanation.
- Response 26 – This response could be construed as thematic on the basis that they could be used together to perform complementary roles in a scenario. However, given the lack of any explicit language that places the two design concepts into an external relation, they cannot be coded as thematic.
- Response 27 – The same rationale applies as for response 26.

Table 7-3 - List of similarity explanations coded as ‘potentially thematic’.

Response #	Response (verbatim)
24	*both related to the office chair*
25	*very similar, use case could be developed in the first to meet the needs that the second has identified.*
26	*one transports one makes it.*
27	*One is a method of collecting the water while the other is a method of transportation*

7.1.2 Discussion of Sim-P2

To probe the number and type of cognitive processes involved in design concept similarity judgements, participants were asked to rate the similarity of pairs of design concepts and explain their ratings. Three research questions were asked: (RQ-S1) do the explanations contain featural or thematic reasons for similarity judgements, (RQ-S2) what is the prevalence of each type of explanation, and (RQ-S3) what is the relationship between the similarity of the pairs and the type of explanation?

Responses to the explanation task were coded as feature-based if they mentioned the common or different features of a pair of design concepts, and thematic if they contained evidence of an external and complementary relationship between the two design concepts. The PDE participants predominantly (93.85%) provided feature-based explanations for their similarity ratings. Of the remaining responses, none could clearly be labelled as feature-based or thematic. This indicated that designers made similarity judgments solely through a feature-based process, i.e., via comparison. The answers to RQs-S1 and S2 rendered the third redundant; only one explanation type was identified and so it was not possible to measure the relationship between similarity and explanation-type.

Two limitations of Sim-P2 are the sample size and the confirmationist reasoning used to provide evidence of designer cognition. 11 participants provided 439 explanations to 40 pairs of design concepts. Prior studies of domain-general similarity judgements have found individual differences in similarity judgements (Simmons and Estes, 2008), where participants tend to consistently favour featural or thematic stimuli in a forced-choice task. It is possible that the 11 participants all favoured featural judgements of similarity ratings, but other participants might favour thematic judgements. Additionally, an absence of evidence for thematic processing does not imply evidence of absence. As thematic relations are thought to have an additive influence on similarity ratings (Section 3.2.3), there might be an unobserved influence of thematic processing that is not dominant enough to come through in the similarity explanations but may nonetheless contribute to a perception of increased similarity. Despite these limitations, the evidence from Sim-P2 was used to justify moving on to test the predictions of the Structural Alignment model. The main implication of these limitations is that the findings are used to provide evidence *for* comparison, but not to rule out thematic processing entirely.

7.2 Similarity experiment 3 (Sim-P3) Testing the predictions of the Structural Alignment model of similarity judgements

The results of the previous experiment indicated that similarity judgements were carried out via feature-based processing and no evidence of thematic processing was found. This was taken as justification for moving on to evaluate predictions of the Structural Alignment model of comparison-based similarity judgements. Five hypotheses were proposed (Section 4.2.2.1, and summarised in Table 7-4, below) that make predictions about the relationships between outcome measures from two tasks: rated similarity (from a similarity rating task) and the number of listed commonalities and differences (from a commonality and difference listing task). The methods and analysis are adopted from Markman and Gentner (1993), and Markman and Gentner (1996). If, as predicted, all hypotheses relating to the Structural Alignment model are supported, it would indicate that the comparison process identified through the results of the previous experiment occurs via a process of structural alignment.

Table 7-4 - Similarity model hypotheses

Ref.	Hypothesis
H – S1a	Similarity should increase as a function of commonalities and decrease as a function of differences and commonalities should influence similarity more than differences.
H – S1b	Alignable differences should be more important in evaluating similarity comparisons than nonalignable differences (Markman and Gentner, 1996).
H – S2	Similar concepts should be associated with an increased number of commonalities and dissimilar concepts should be associated with a decreased number of commonalities.
H – S3	There should be a numerical link between commonalities and alignable differences.
H – S4	Alignable differences should be more numerous than nonalignable differences.

7.2.1 Analytic considerations and definitions of variables

7.2.1.1 Selection of statistical tests

The statistical tests used were multiple regression analyses (Laerd Statistics, 2015b) for H1a and H1b, Pearson's product-moment correlations (Laerd Statistics, 2018) for H2 and H3, and the Wilcoxon signed-rank test of group differences (Laerd Statistics, 2015c) for H4. In addition to the general considerations for selection of statistical tests (Section 5.1.5), additional considerations were made based on findings from prior research.

Regression analyses were used with the caveat that prior research had highlighted issues with multicollinearity. Markman and Gentner (1993) carried out regression analyses with similarity as a criterion variable and Com, AD and ND as predictor variables; the same analysis planned for H1b in this thesis. They found that Com was a positive independent predictor and ND was a significant negative predictor, but contrary to expectations, AD did not significantly predict similarity. They proposed that this may be attributable to the significant positive correlation between commonalities and alignable differences. This is predicted by the Structural Alignment model as shown in Figure 3-3. They appear to have highlighted an issue of multicollinearity, noting that "a linear model cannot separate the (positive) impact of commonalities on similarity from the (predicted negative) impact of alignable differences on similarity" (Markman and Gentner, 1996, p.238). To further test this prediction, they conducted an additional experiment using a forced-choice task and demonstrated that alignable differences did indeed influence similarity more than nonalignable differences (Markman and Gentner, 1996, Exp. 2) (but see Estes and Hasson (2004) for empirical concerns). In the present research, regression analyses are carried out, checking for multicollinearity, with the acknowledgement that the same issue might arise and require additional empirical research.

Correlational analyses were used to analyse the relationship between rated similarity and the number of commonalities (H2) and the relationship between commonalities and alignable differences (H3). An alternative approach would have been bin the data into groups of e.g., low, medium and high similarity and conducting statistical tests of group differences) (e.g., (Markman and Gentner, 1993b, 1996)). This is beneficial in that it allows the use of individual-level data. Correlations require the use of mean or median values for each of the 30 pairs of design concepts because each similarity rating has a corresponding list of commonalities *or* differences, but not both. This means that inferences about the stimuli can only be made at the aggregate level. For example, in investigating the relationships between rated similarity and the number of listed commonalities, it may be possible to say that, on average, highly similar concepts have many commonalities. It would not be possible to conclude that all participants that rate a pair of concepts as highly similar would list many commonalities. However, binning the data loses a valuable quality of the data; that the central tendencies of the stimuli vary fairly linearly across the similarity scale (Section 6.4.3). Moreover, decisions about how many bins to create would be arbitrary and may have unforeseeable influences on the analyses. Acknowledging the aforementioned limitations, correlational analyses are used in the subsequent analyses to preserve the variation in base-concept similarity (Section 6.4.3). Specifically, Pearson's product moment correlations were used to be consistent with prior research (Markman and Gentner, 1996).

7.2.1.2 Definition of variables

The variables used in the analyses of the similarity hypotheses are defined in Table 7-5. Three variables are used in analyses involving similarity ratings. This is because participants provided similarity ratings for every pair, but only listed the commonalities *or* the differences for any given pair. To limit the potential influence of group differences it is beneficial to correlate responses provided by the same participants. For example, when investigating the relationship between similarity and listed commonalties, it would be beneficial to use only similarity ratings with matching commonality ratings from the same participant. To this end, three measures of similarity are used: Total similarity 'Sim(all)', Similarity matched with commonalities 'Sim(com)' and Similarity matched with differences 'Sim(dif)'. The three similarity variables are tested in Section 7.2.3.3 to assess the validity of using Sim(com) and Sim(dif) as proxies for Sim(all).

Table 7-5 - Shorthand designations and definitions for variables in Similarity Experiment 3.

Variable	Definition
Sim(all)	Rated similarity derived from all similarity responses. Sim(all) is used when investigating the relationships between rated similarity, number of listed commonalities *and* number of listed differences.
Sim(com)	Rated similarity from responses with matching commonalities. Sim(com) is used when investigating the relationships between rated similarity and the number of listed commonalities.
Sim(dif)	Rated similarity from responses with matching difference responses. Sim(dif) is used when investigating the relationships between rated similarity and the number of listed differences
Com	Number of listed commonalities
Dif(tot)	The number of listed total differences. The sum of AD and ND
AD	The number of listed alignable differences.
ND	The number of listed nonalignable differences.

7.2.2 Descriptive statistics and example responses

37 individuals participated in Similarity Experiment 3, rating the similarity of 30 pairs of design concepts and listing either the commonalities or differences for that pair.

7.2.2.1 Example responses

An example of the commonalities and differences listed for a pair of design concepts is shown below in Figure 7-2 and Table 7-6. Differences have been coded as alignable or nonalignable. Alignable differences being those that vary along a common dimension, and nonalignable differences being those that refer to a feature of one design concept but not the other.

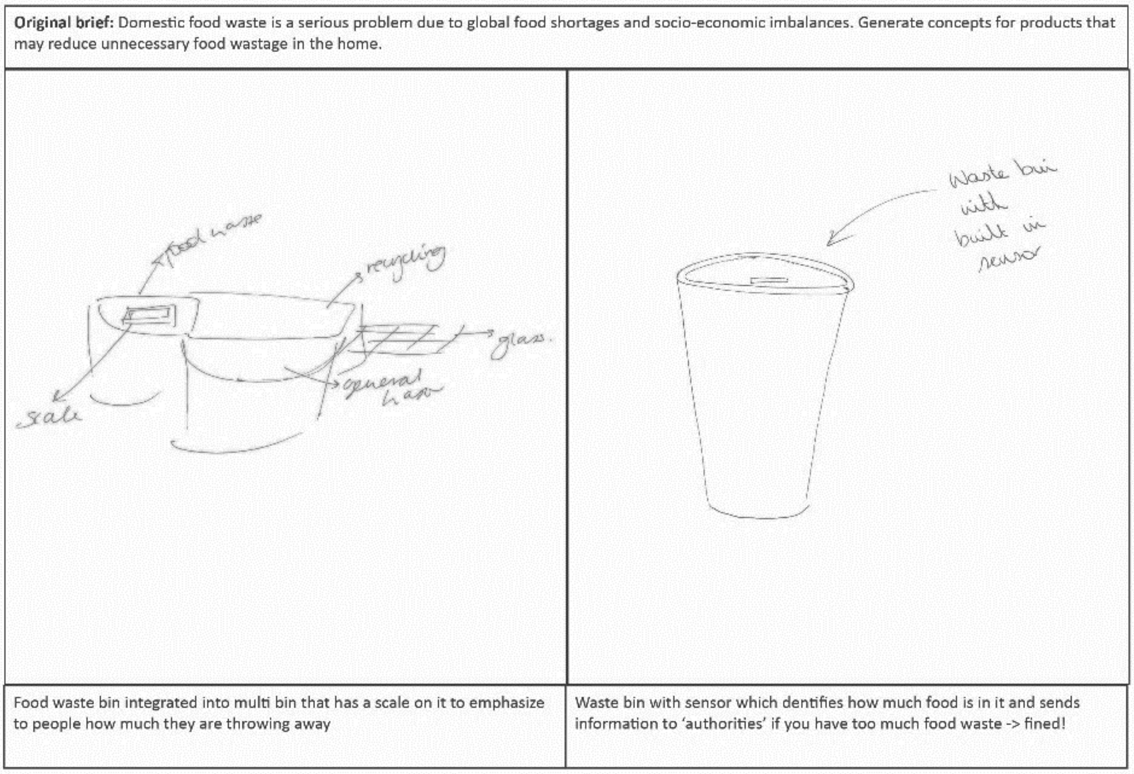


Figure 7-2 – Pair of design concepts Sc_06.

Table 7-6 – Examples of commonalities and differences listed for stimulus Sc06.

ID	Commonalities	ID	Differences	Code
2	• Both uses sensors to measure food waste • both records volume of food • both notifies someone of waste • both are bin like structures • both require waste to be inserted	3	One focuses on punishment as a deterrent, the other provides information	AD
			One is linked to a mobile network, the other is completely local	AD
			One product is integrated alongside a conventional system, the other is integrated into a conventional system	AD
			One utilises feedback to the user, the other only sends feedback and information to a third-party (authorities)	AD
6	Both products use the concept of telling the person just how much they are wasting by using some sort of scale and limit to consumption	5	one gives fines	ND
			one identifies each food	ND
			one alerts more than just the user to the waste produced	AD
10	• Both sense food amount. • Both discourage food waste. • Both give out a measurement. • Both are bins.	11	Right can be a compact design vs left requiring different areas.	AD
			left uses different materials such as glass etc vs Right can use one main material.	AD
12	• both identify how much food is inside by weight • both alert user to how much they are affecting the environment. • both are variants of household bins	15	one notifies the user the other fines them	AD
			one is attached onto existing bins	ND
			one encourages good practice	ND
			one is deterring to users	ND

ID = participant number, AD = alignable difference, ND = nonalignable difference

7.2.2.2 Responses, exclusions and outlier removal

Two participants were excluded and some did not complete the entire listing task. The two exclusions were based on poor quality responses to the listing task. Table 7-7 shows a representative example from each participant. Participant 16 simply described both concepts in every pair. Participant 31 predominantly listed the features of the two concepts, and it was unclear whether any of the features were supposed to be linked. Some participants exercised their right to leave the experiment at any point. This resulted in fewer responses to the listing task but the randomisation of the stimuli spread this loss across the sample. The mean number of responses for each variable, after excluding two participants and before outlier removal, is n = 31 for Sim(all) and n = 15 for all other variables. Full details of the number final number of responses are provided in Section 7.2.2.3 after outlier removal.

Table 7-7 - Representative example responses for the two participants excluded from analyses

ID	Example response
16	one is an app on a fridge interface telling the user what is going out of date by tracking the sell date. the other is a fridge that gives a signal of when something is going out of date but doesn't tell the user what is going out of date
31	L: pocket sized R: sizeable unit R: vacuum pump not integrated into 1 unit r: larger portions

ID = participant number.

Outliers were identified from boxplots and responses 3 or more units from the edge of the interquartile range (IQR) were removed. Outliers were removed from similarity ratings as these could have arisen from data-entry errors or lapses in the participant's attention. No values for the number of listed commonalities or differences were considered outliers as there is no conceptual or methodological reason to discount e.g., an unusually large number of responses to the listing task.

Figure 7-3 shows boxplots for the 1049 values of Sim(all) across 30 pairs of design concepts and outliers have been marked with unique numbers that are used for reference can be considered arbitrary. Using the rule of 3 units or more from the edge of the inter-quartile range (IQR) provided a less severe removal of outliers than the default option in

SPSS. For example, SPSS flags data point #691 (the Asterisk above pair 19, Figure 7-3) as an outlier, but a rating of 2 is not problematic for a set of values Mdn = 1. 20 responses were removed using this method. Boxplots for similarity ratings after manual outlier removal are shown in Figure 7-4.

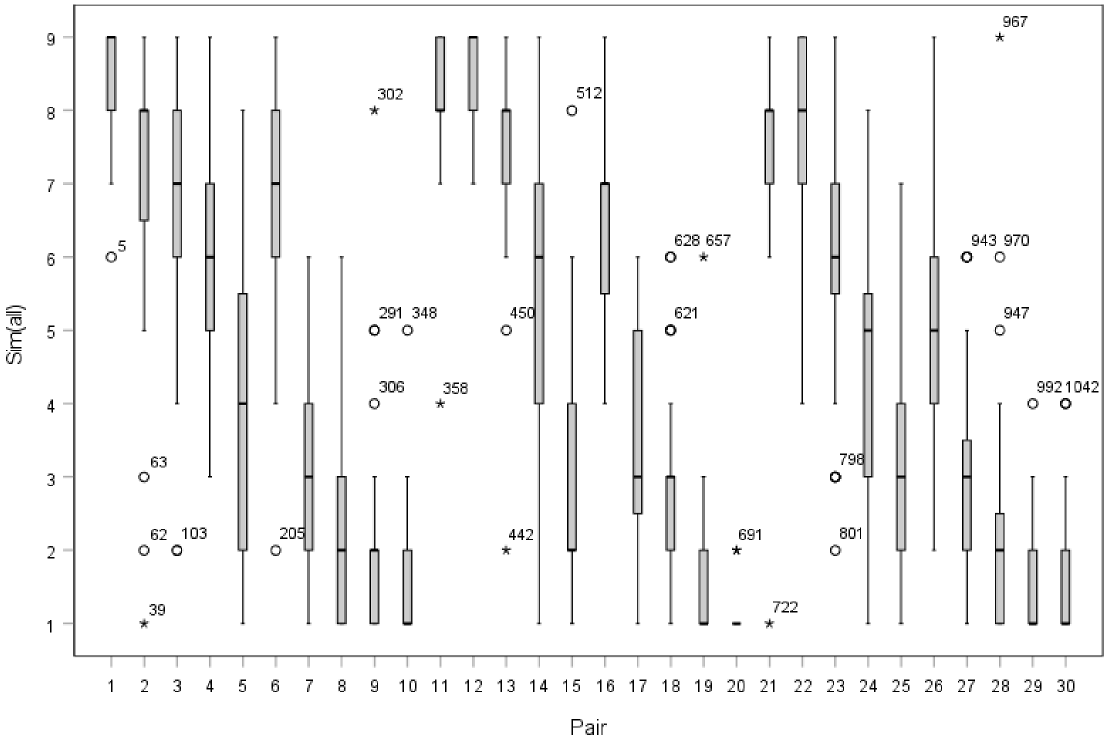


Figure 7-3 – Boxplot of Sim(all) for each pair of design concepts before outlier removal. Outliers (circles and asterisks) shown here are calculated by SPSS and numbers can be considered arbitrary.

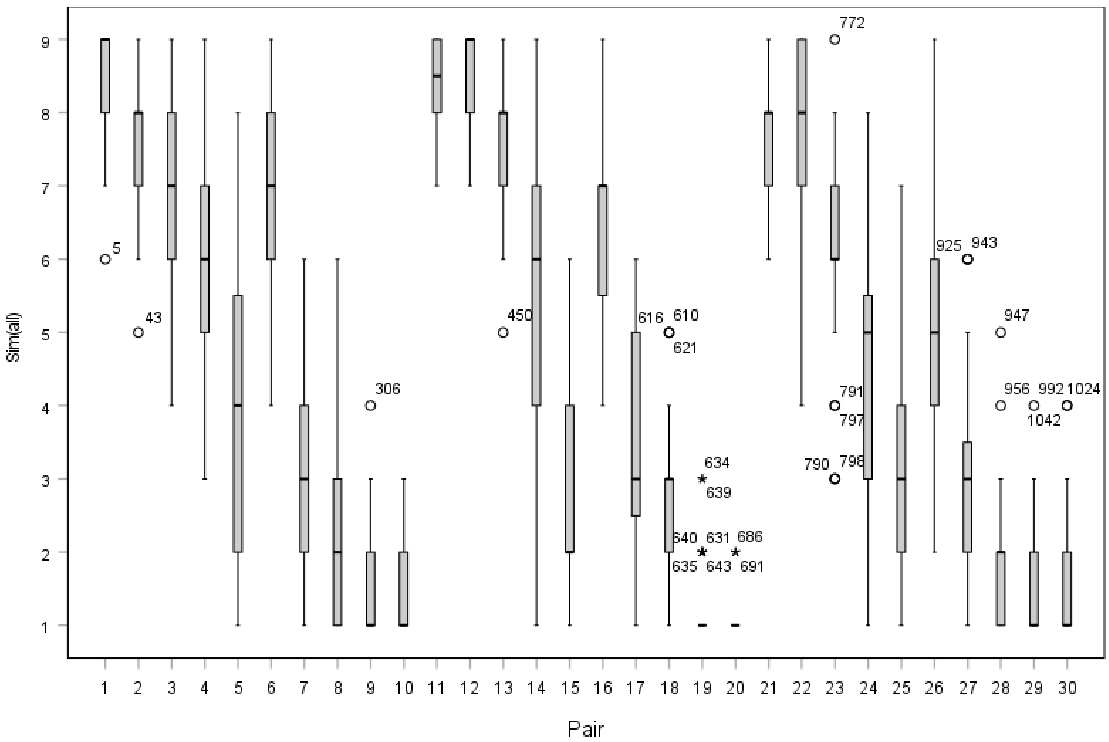


Figure 7-4 **–** Boxplot of Sim(all) for each pair of design concepts after outlier removal. Outliers (circles and asterisks) shown here are calculated by SPSS after manual outlier removal but are not treated as outliers for subsequent analyses.

7.2.2.3 Data summary

Table 7-8 presents a summary of the mean and median values, standard deviations, the number of responses and the total number of items listed for each variable across the three design tasks after outliers have been removed. The same values per stimulus can be found for each pair of design concepts in Table A7-4 to Table A7-6.

Table 7-8 Values for all variables across each of the three design tasks and in total.

DT		Sim(all)	Sim(com)	Sim(dif)	Com	Dif(tot)	AD	ND
	M	4.81	4.84	4.83	3.28	3.31	2.25	1.01
03	*SD*	*2.77*	*2.79*	*2.85*	*1.84*	*1.82*	*1.43*	*1.33*
(pairs	Sum	---	---	---	489	507	345	155
1-10)	Mdn	5.00	5.00	5.00	3.00	3.00	2.00	1.00
	n	340	149	153	149	153	153	153
	M	4.82	4.75	4.92	3.28	3.34	1.83	1.49
06	*SD*	*3.00*	*2.99*	*3.00*	*1.80*	*1.99*	*1.32*	*1.64*
(pairs	Sum	---	---	---	466	481	264	214
11-20)	Mdn	5.00	5.00	5.00	3.00	3.00	2.00	1.00
	n	343	142	144	142	144	144	144
	M	4.21	4.25	4.16	3.52	3.97	2.70	1.29
07	*SD*	*2.59*	*2.61*	*2.54*	*1.93*	*2.28*	*1.65*	*1.74*
(pairs	Sum	---	---	---	553	612	416	199
21-30)	Mdn	4.00	4.00	4.00	3.00	3.00	3.00	1.00
	n	346	157	154	157	154	154	154
	M	4.61	4.61	4.63	3.37	3.55	2.27	1.26
	SD	*2.80*	*2.80*	*2.81*	*1.86*	*2.06*	*1.51*	*1.59*
Total	Sum	---	---	---	1508	1600	1025	568
	Mdn	4.00	4.00	4.00	3.00	3.00	2.00	1.00
	n	1029	448	451	448	451	451	451

Note: DT = Design task, M = Mean, Mdn = Median, SD = Standard Deviation, n = number of responses, Sum = total number of items listed across all responses.

7.2.3 Pre-analysis data checking

Two sets of analyses are conducted before the main analysis:

1. It is beneficial to analyse the responses to all 30 pairs as one homogeneous set to maximise the available data points for performing correlation analyses. This assumes some degree of homogeneity across the three design tasks. To test this assumption, the median responses for each outcome measure are compared across the three design tasks are compared (Sections 7.2.3.1 and 7.2.3.2). Significant differences in the response distributions between design tasks may suggest that the design task itself is an extraneous variable, which may in turn influence the analysis or interpretation of the data.
2. Sim(all) has been split into two proxy measures, Sim(com) and Sim(dif), to satisfy methodological requirements (Section 7.2.1). A strong correlation with a high

percentage of the variance explained would indicate that Sim(com) and Sim(dif) variables are suitable proxy measures for Sim(all). Thus, the degree of association between mean Sim(com) and mean Sim(dif) is examined (Section 7.2.2.3).

The implications of the results from these tests are discussed in Section 7.2.3.4.

7.2.3.1 Analysis of rated similarity across design tasks

Individual-level data are used to test for differences in the similarity responses (Sim(tot), Sim(com) and Sim(dif)) to the stimuli from the three design tasks (DT03, DT06 and DT07). None of the similarity responses were normally distributed as assessed by Shapiro-Wilk's test, ($p < .001$) and as is supported visually in Figure 7-5. It should be expected that rated similarity is not normally distributed as the stimuli set has been shown to span a range of similarity values along the 1-9 scale (Section 6.4.3).

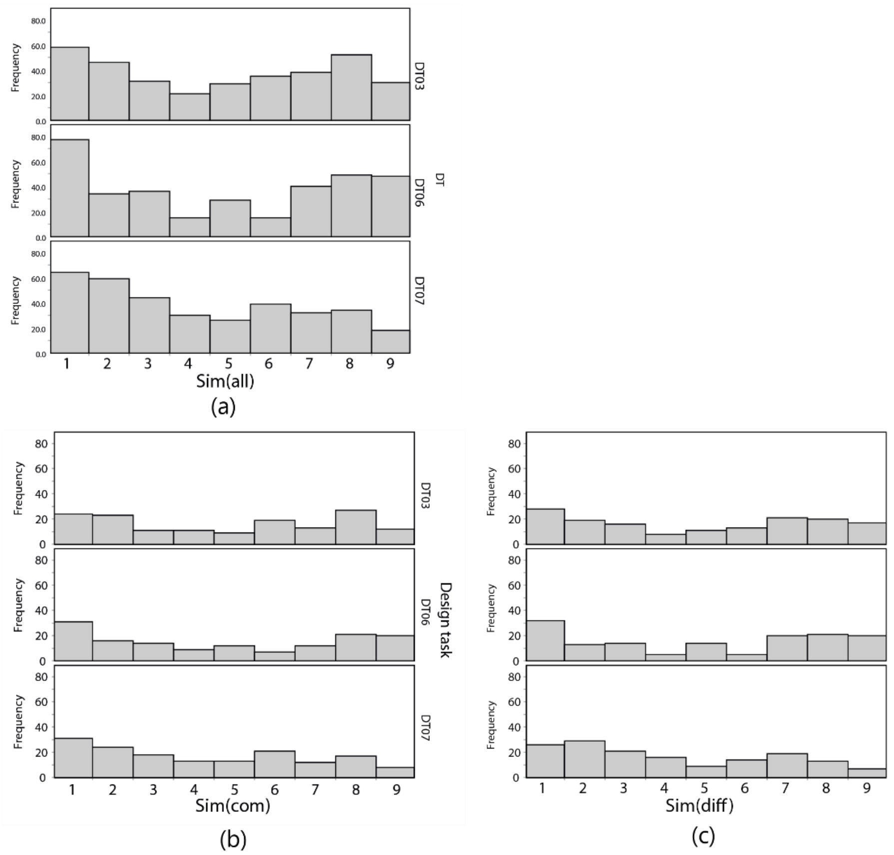


Figure 7-5 – Frequency histograms for a) Sim(all), b) Sim(com) and c) Sim(dif).

Due to the violation of normality, the non-parametric Kruskal-Wallis H test was used. An assumption of the Kruskal-Wallis H test is that the responses for the three design tasks have the same distribution shape. It can be seen in Figure 7-5 that the shape of the distribution of similarity responses is similar for Sim(all), Sim(com) and Sim(dif). Thus, the Kruskal-Wallis H test was used to compare the median value of the responses across three design tasks (Figure 7-6). Specifically, this involves taking all responses to a given design task, i.e. responses by 30 participants for 10 concepts in each task, and comparing them with the responses for the other tasks.

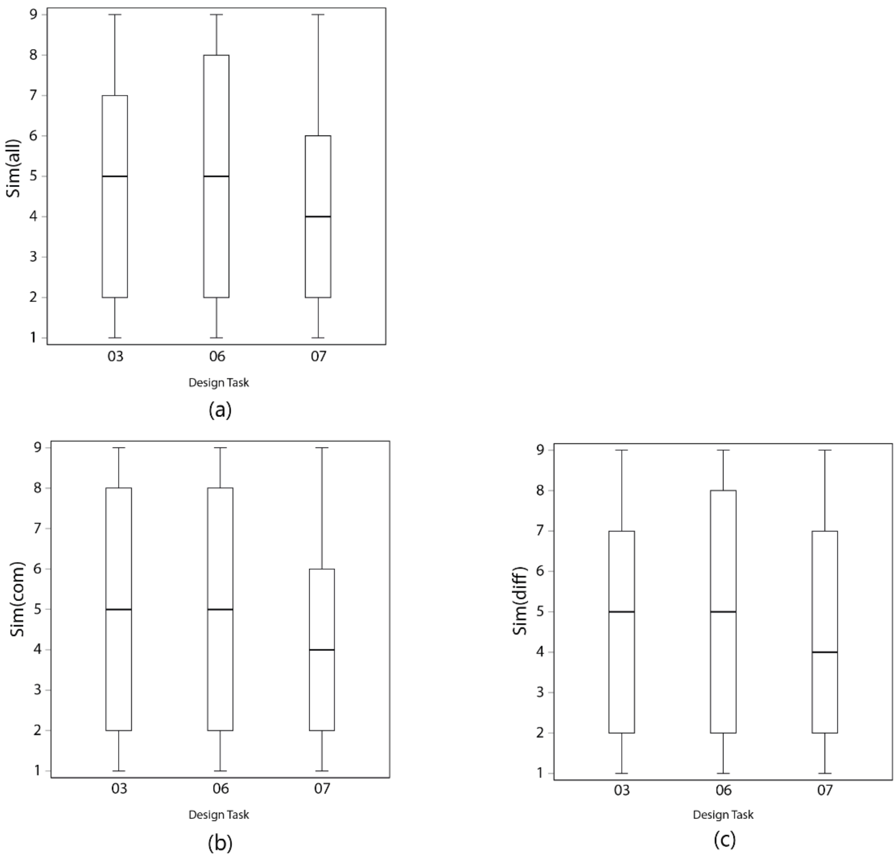


Figure 7-6 - Boxplots for three measures of rated similarity for each design task.

Differences across design tasks for Sim(all). Median similarity ratings were significantly different between design tasks, χ^2 (2) = 9.266, *p* = .010. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median rated similarity between DT07 and DT03 (*p* = .025), and DT07 and DT06 (*p* =.026), but not for DT06 and DT03 (*p* = 1.000).

Differences across design tasks for Sim(com). Median rated similarity changed between DT03 DT06 and DT07, but the differences were not statistically significant χ^2 (2) = 3.668, *p* = .160.

Differences across design tasks for Sim(dif). Median rated similarity changed between DT03, DT06 and DT07, but the differences were not statistically significant χ^2 (2) = 5.552, *p* = .062.

In summary, the median similarity responses to the 10 pairs of design concepts in DT07 were significantly lower than those in DT06 and DT03 when using the variable Sim(all), but there was no significant difference when using the proxy measures Sim(com) and Sim(dif). The implications of this are discussed in Section 7.2.3.4.

7.2.3.2 Analysis of the number of listed items across design tasks

A similar procedure to the one used in the previous section is used to test for differences in the outcome measures for the listing task (Com, Dif(tot), AD, ND) across the three design tasks (DT03, DT06, DT07) (Figure 7-7). Again, individual-level data are used. None of the listing task variables were normally distributed, as assessed by Shapiro-Wilk's test, ($p < .001$) and so the non-parametric Kruskal-Wallis test was used again.

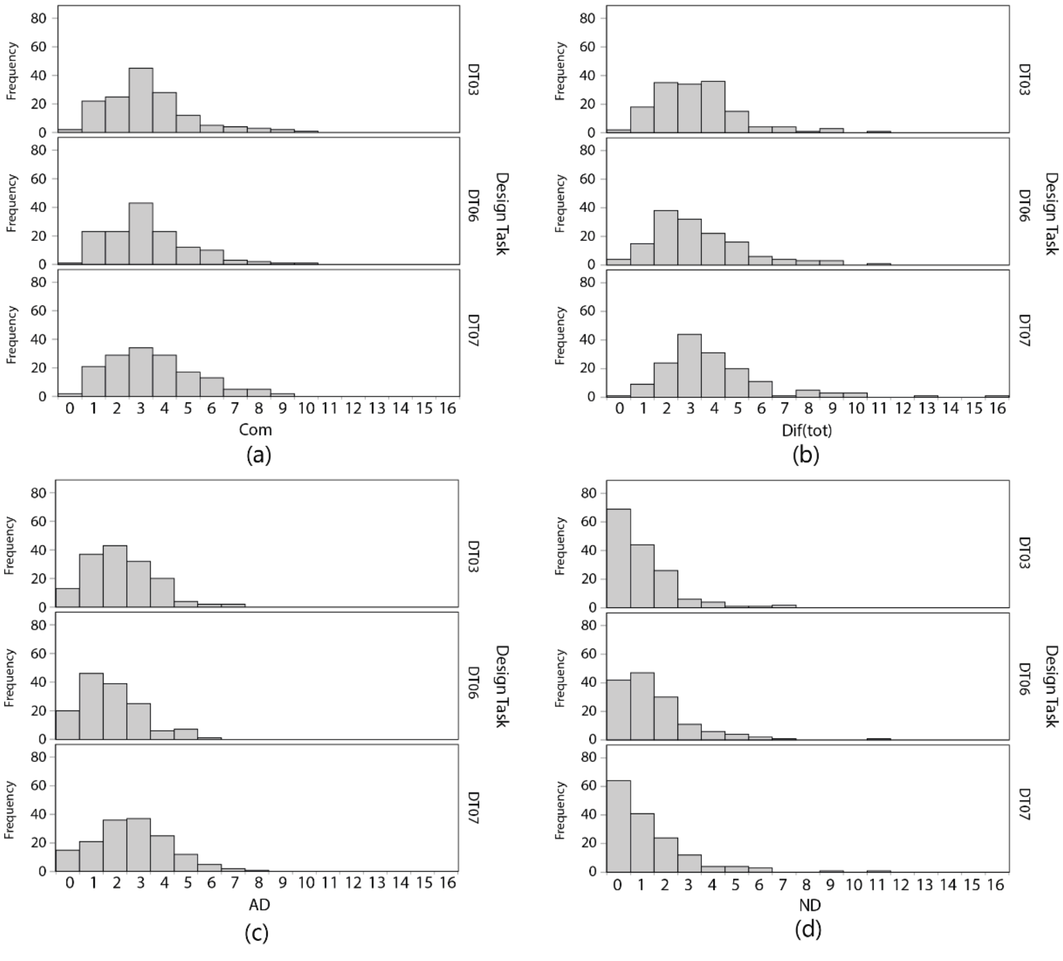


Figure 7-7 – Frequency histograms for the number of listed a) Com, b) Dif(tot), C) AD, and d) ND.

The null hypothesis for the Kruskal-Wallis H test is that the distributions of the groups are equal. If it can also be assumed that the *shape* of the distribution is the same for the groups, then the differences between groups can be attributed to a difference in medians. Through visual inspection of the histograms (Figure 7-7) and boxplots (Figure 7-8), the distributions of each outcome measure across the three design tasks appear somewhat similar, indicating that it may be appropriate to compare the group medians. However, in the case of the NDs, a Kruskal-Wallis H test shows that there are significant differences across the three design tasks, even though the median value is the same across the three design tasks (Mdn = 1, see Figure 7-8e). Thus, for NDs, the Kruskal-Wallis H test cannot be used to compare differences in the median values. It can, however, be used to compare the *mean ranks* of the distributions, although this has comparatively less descriptive power. Thus, for consistency in the subsequent statistical tests, the Kruskal-Wallis H test is used as a comparison of mean ranks for the number of listed items (C, Dif(tot), AD, ND) in each design task. All post hoc analyses are pairwise comparisons that were performed using Dunn's (Dunn, 1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented.

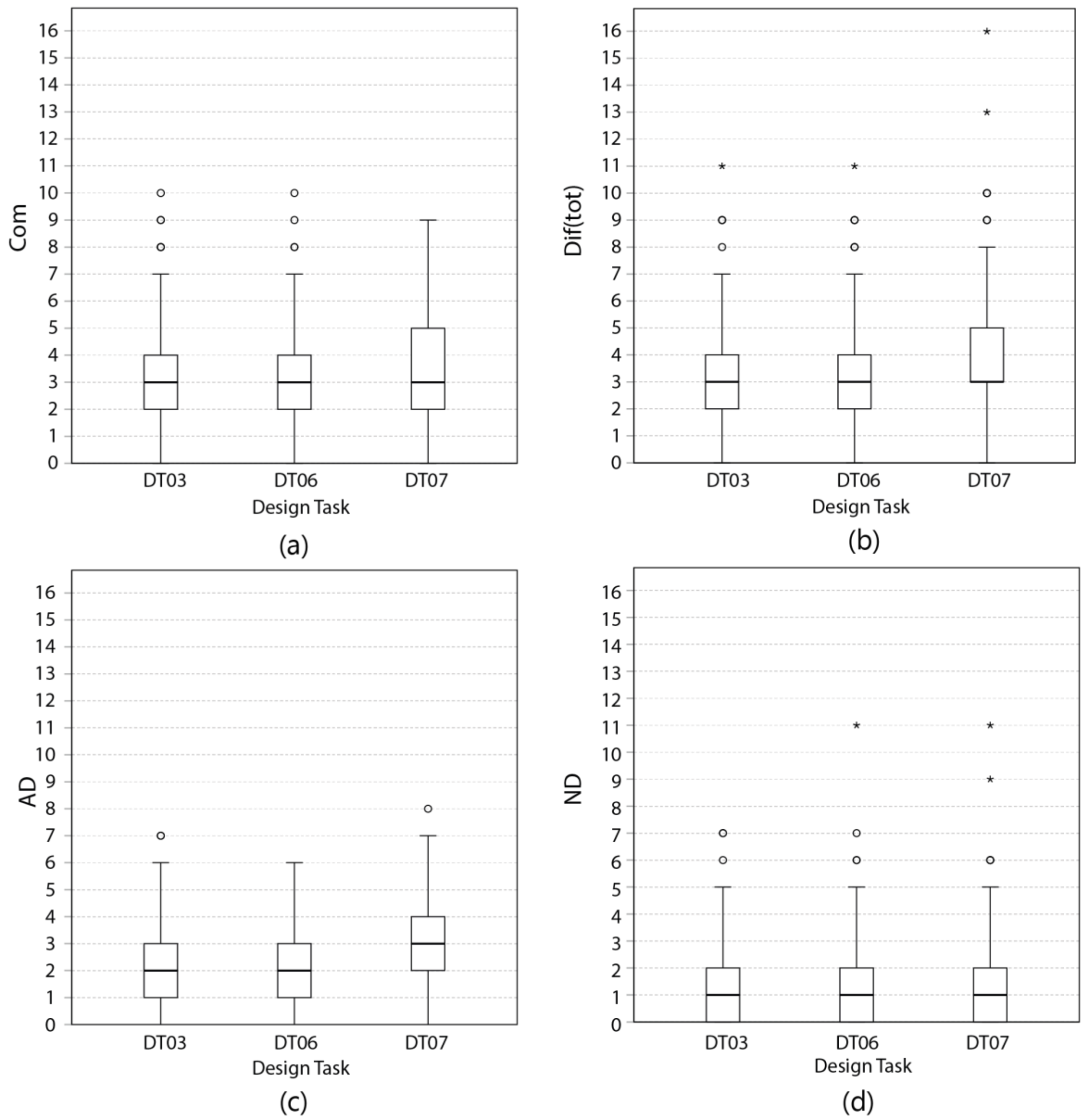


Figure 7-8 – boxplots for the number of items listed in the commonality and difference listing task for three design tasks. Com = commonalities, Dif(tot) = total listed differences, AD = alignable differences, ND = nonalignable differences.

Number of listed commonalities across three design tasks. The mean rank of number of listed commonalities changed from DT03 (mean rank = 218.46) to DT06 (mean rank = 219.13) to DT07 (mean rank = 235.09), but the differences were not statistically significant, χ^2 (2) = 1.677, p = .432.

Number of listed total differences across three design tasks. The number of listed total differences were statistically significantly different between the design tasks, χ^2 (2) = 9.845, p = .007. The post hoc analysis revealed statistically significant differences in rated similarity between DT03 (mean rank = 213.85) and DT07 (mean rank = 252.22) (p = .026), and DT06 (mean rank = 210.87) and DT07 (p = .016), but not for DT03 and DT06 (p = 1.000).

Number of listed alignable differences across three design tasks. The number of listed alignable differences were significantly statistically different between design tasks χ^2 (2) = 24.723, p < .001. The post hoc analysis revealed statistically significant differences in number of alignable differences between DT03 (mean rank = 226.02) and DT06 (mean rank = 187.99) (p = .031), DT06 and DT07 (mean rank = 261.52) (p < .001), and DT03 and DT07 (p = .044).

Number of listed nonalignable differences in each design task. The number of listed nonalignable differences was significantly statistically different between design tasks χ^2 (2) = 9.033, p = .011. The post hoc analysis revealed statistically significant differences in number of nonalignable differences between DT03 (mean rank = 206.60) and DT06 (mean rank = 249.70) (p = .009) but not for any other pairs of values, including DT07 (mean rank = 223.11).

In summary, there were statistically significant differences in the distribution of the number of listed differences across design tasks (including Diff(tot), AD and ND), but not for the distribution of the number of listed commonalties.

7.2.3.3 Test of association between Sim(com) and Sim(dif)

A Pearson's product-moment correlation was run to assess the relationship between mean Sim(com) and mean Sim(dif) for 30 pairs of design concepts. Neither mean Sim(com) nor mean Sim(dif) were normally distributed, as assessed by Shapiro-Wilk's test, Sim(com) (p = *.12),* Sim(dif) (p = *.12*). A Pearson's correlation was carried out despite the violation of normality. Visual inspection of the scatterplot of mean Sim(com) against mean Sim(dif) (Figure 7-9) indicates that there is a linear relationship between the two variables and no outliers. There was a statistically significant, strong positive correlation between Sim(com) and Sim(dif), r(28) = .997, $p < .001$, with 99.4% of the variance explained. The strong positive correlation between the two similarity ratings suggests that the participants responded to the similarity rating task in a similar way, regardless of their allocation in the listing task.

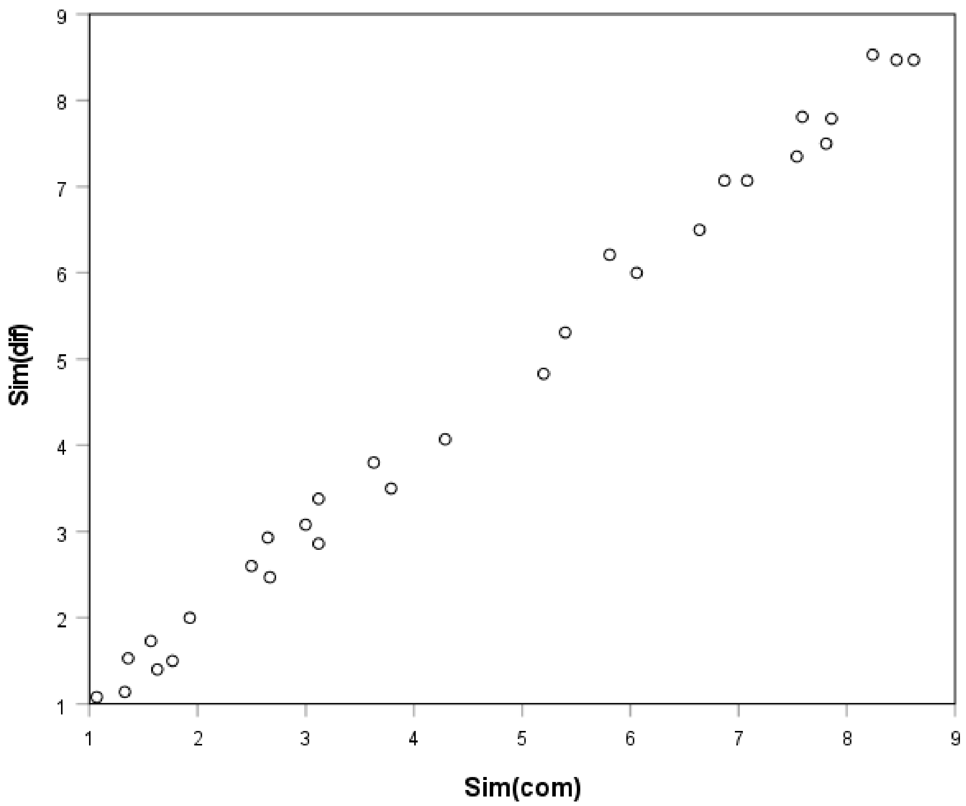


Figure 7-9 – Scatter plot of mean rated similarity with matching commonality rating 'Sim(com)', against mean rated similarity with matching difference rating 'Sim(dif)'. Each point represents one of 30 pairs of design concepts.

7.2.3.4 Implications of data checking findings on further analysis

Two sets of analyses were conducted to determine (i) whether participants responded in a similar way to the design concepts from the three design tasks, and (ii) whether participants rated similarity in the same way regardless of whether they listed the commonalities or the differences of concepts.

With regards to (ii), the use of Sim(all), Sim(com) and Sim(dif) (see Section 7.2.1 for the rationale behind these variables), it appears that participants provided similarity ratings in a homogeneous way regardless of whether they listed commonalities or differences for a given pair of concepts. This should be expected since the similarity ratings occurred before the participants received their listing task instructions. This supports the use of the Sim(com) and Sim(dif) variables as proxy measures for Sim(all).

With regards to (i), significant statistical differences were found for the distribution of mean rated similarity and the distribution of the number of listed differences across the three design tasks. Generally, stimuli in DT07 tended to elicit lower similarity ratings and more listed differences, the number of listed Coms was consistent, but the number of listed ADs and NDs was generally inconsistent across the design tasks. None of these statistical differences violates the requirements of the stimuli, i.e., that they span a range of similarity ratings. Nor are they expected to interfere with any of the correlational or regression analyses. However, some possible implications were identified concerning H4 which prompted additional analyses as reported in that section (Section 7.2.4.4).

7.2.4 Results

For all statistical tests presented in this section (7.2.4), all tests of normality were assessed by Shapiro-Wilk's test unless otherwise stated. Correlational analyses were conducted in some cases despite violations of normality.

7.2.4.1 Testing hypotheses 1a and b – Similarity, commonalities and differences

S-H1a states that *similarity should increase as a function of commonalities and decrease as a function of differences, and commonalities should influence similarity more than differences*. H1a is a prediction common to both the Contrast and Structural Alignment models and was included as a first-pass test of the applicability of featural models generally.

A multiple regression analysis was run to predict the value of mean rated similarity (Sim(all)) based on the mean number of listed commonalities (Com) and the mean number of listed total differences (Dif(tot)). There was linearity as assessed by partial regression plots (Figure 7-10, a and b). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There was one instance in which the studentized deleted residual was greater than +3 standard deviations (value = 3.17). This entry was examined, but there were no apparent data entry or conceptual issues that warranted its removal. There were no leverage values greater than 0.2 and no values for Cook's distance above 1. The assumption of normality was met, as assessed by a P-P Plot (Figure 7-10, c).

The multiple regression model was significant, R^2=0.770, $F(2, 27)$ = 45.107, $p < .001$, with an adjusted R^2 of 0.753. Both commonalities (β=1.598, p <.001) and total differences (β=-.883, p =.006) were significant predictors of Sim(all), explaining 77% of the variance in rated similarity. Regression coefficients and standard errors can be found in Table 7-9. That the unstandardized regression coefficient is greater for commonalities than differences indicates that commonalities count more towards similarity than differences count against similarity, consistent with previous findings in non-design contexts (Krumhansl, 1978; Markman and Gentner, 1993b, 1996; Tversky, 1977).

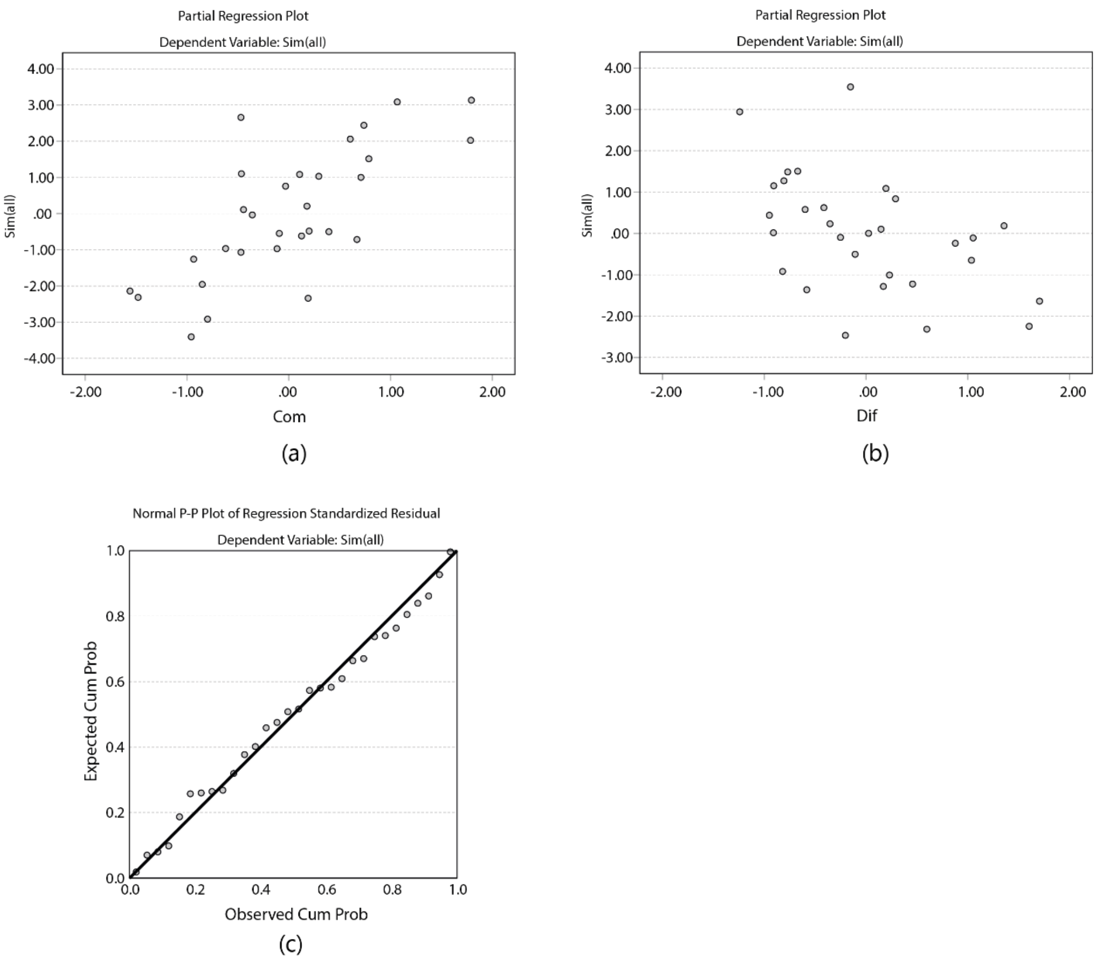


Figure 7-10 – Charts for multiple regression to determine how much Sim(all) changes with Com and Dif(tot). Panels show a) Partial regression plot of Sim(all) and Com, b) partial regression plot of Sim(all) and Dif(tot), and c) P-P plot of standardized residuals.

Table 7-9 - Summary of multiple regression analysis for Hypothesis 1a.

Variable	***B***	**95% CI for B**		***SE_B***	***Beta***	***Sig.***
		LL	UL			
Intercept	2.398	-1.305	6.101	1.805		
Com	1.598	1.008	2.188	.288	.633	<.001
Dif(tot)	-.883	-1.488	-0.278	.295	-.341	.006

Note. *B = unstandardised regression coefficient; SEb = Standard error of the coefficient; Beta = standardized coefficient; CI = Confidence Interval; LL = Lower limit; UL = upper limit.*

S-H1b states that *alignable differences should be more important in the judgement of similarity than nonalignable differences.* That is, changes in the number of listed alignable and nonalignable differences should predict changes in similarity, and the unstandardized regression coefficient for alignable differences should be greater than that of the nonalignable differences. A multiple regression was run to determine whether changes in the independent variables Com, AD and ND, predict changes in the dependent variable Sim(all).

There was linearity as assessed by partial regression plots (Figure 7-11, a-c). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. The studentized deleted residual for pair 3 was greater than +3 standard deviations (value = 3.48). The leverage values were greater than 0.2 for four pairs (pairs 16, 20, 21, 26). There were no values for Cook's distance above 1. There were no data entry or conceptual issues that could be attributed to the influential points and thus no action was taken. The assumption of normality was met, as assessed by a P-P plot (Figure 7-11, d).

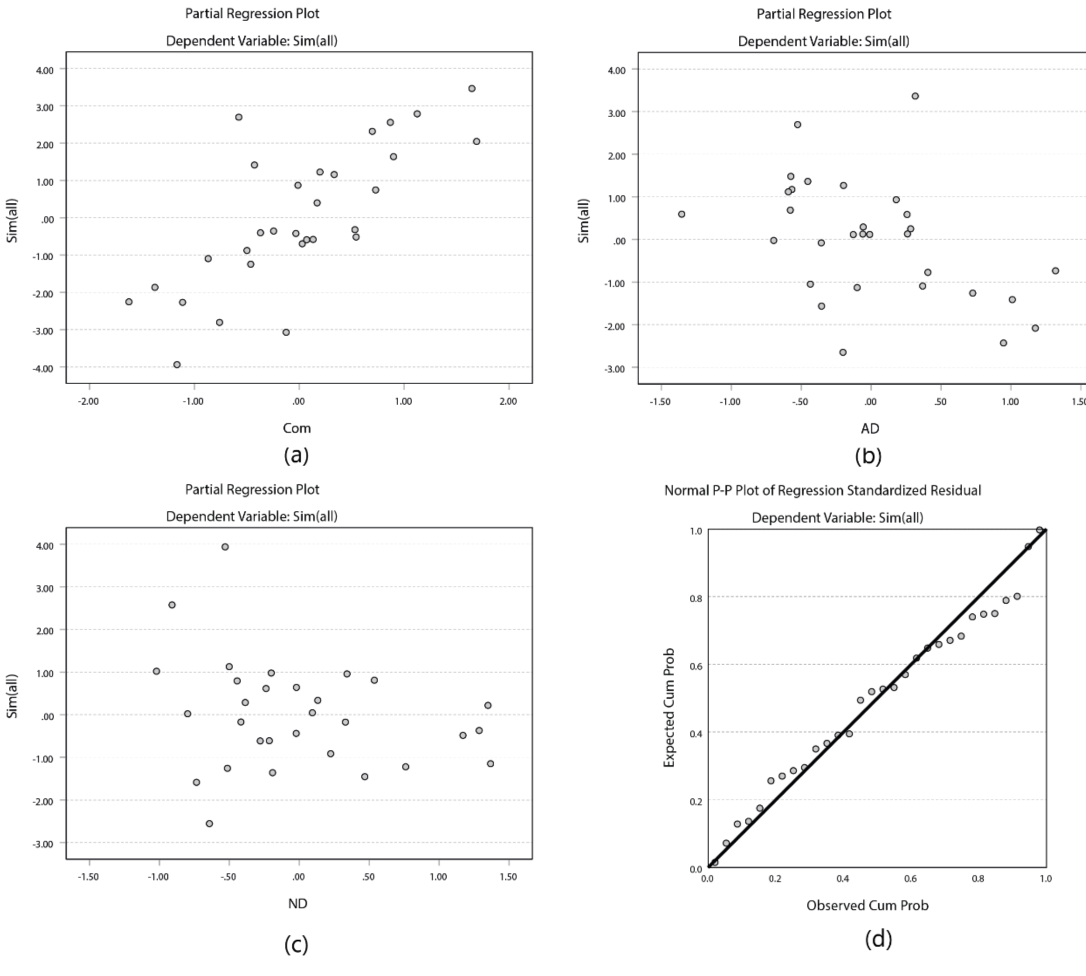


Figure 7-11 - Charts for multiple regression to determine how much Sim(all) changes with Com, AD, and ND, a) Partial regression plot of Sim(all) and Com, b) partial regression plot of Sim(all) and AD, c) m partial regression plot of Sim(all) and ND, and d) P-P plot of standardized residuals.

Table 7-10 - Summary of multiple regression analysis for hypothesis 3.

Variable	*B*	*95% CI for B*		SE_B	*Beta*	*Sig.*
		LL	*UL*			
Intercept	1.872	-1.967	5.711	1.868		
Com	1.695	1.075	2.315	.302	.676	.000
AD	-1.025	-1.867	-.183	.410	-.262	.019
ND	-.480	-1.245	.286	.372	-.148	.209

Note. *B* = unstandardised regression coefficient;, SE_a = Standard error of the coefficient; Beta = standardised coefficient, CI = Confidence Interval, LL = Lower limit, UL = upper limit.

The multiple regression model was significant, R^2=0.761, $F(3, 26) = 26.092$, $p < .001$, summary in Table 7-10, with an adjusted R^2 of 0.733. Commonalities (β=.1.722, p <.001) and alignable differences (β=-1.036, p =.016) added statistically significantly to the variance in the similarity ratings. The regression coefficient for nonalignable differences was not significant (β=-0.469, p =.210). In other words, mean rated similarity increases by 1.722 for every additional unit of mean listed commonalities when all other variables are held the same and decreases by 1.036 for every alignable difference listed when all other variables are held the same.

The result of the first regression analysis indicates that similarity changes as a positive function of commonalities and a negative function of total differences and commonalities contribute more to similarity than differences detract from it. This is consistent with both the Contrast and Structural Alignment models. The results of the second regression analysis show that it is alignable differences, not nonalignable differences, that are responsible for the detraction in perceived similarity.

7.2.4.2 Testing hypothesis 2 – Correlational relationship between similarity and the number of commonalities

S-H2 states that *similar concepts should be associated with an increased number of commonalities and dissimilar concepts should be associated with a decreased number of commonalities*. A Pearson's product-moment correlation was run to assess the relationship between mean Sim(com) and mean Com for all 30 pairs of design concepts. Visual inspection of the scatterplot of Com against Sim(com) (Figure 7-12) indicates that there is a linear relationship between the two variables and no outliers. The mean values for Com were normally distributed, (p = .101), but the mean values for Sim(com) were not (p = 0.12).

There was a statistically significant, strong positive correlation between mean rated similarity (Sim(com)) and mean number of listed commonalities (Com), $r(28) = .827$, $p < .001$ (one-tailed), with Sim(com) explaining 68.4% of the variance in Com. This shows that on average, concepts with higher similarity ratings have more listed commonalities, and concepts with lower similarity ratings have fewer listed commonalities.

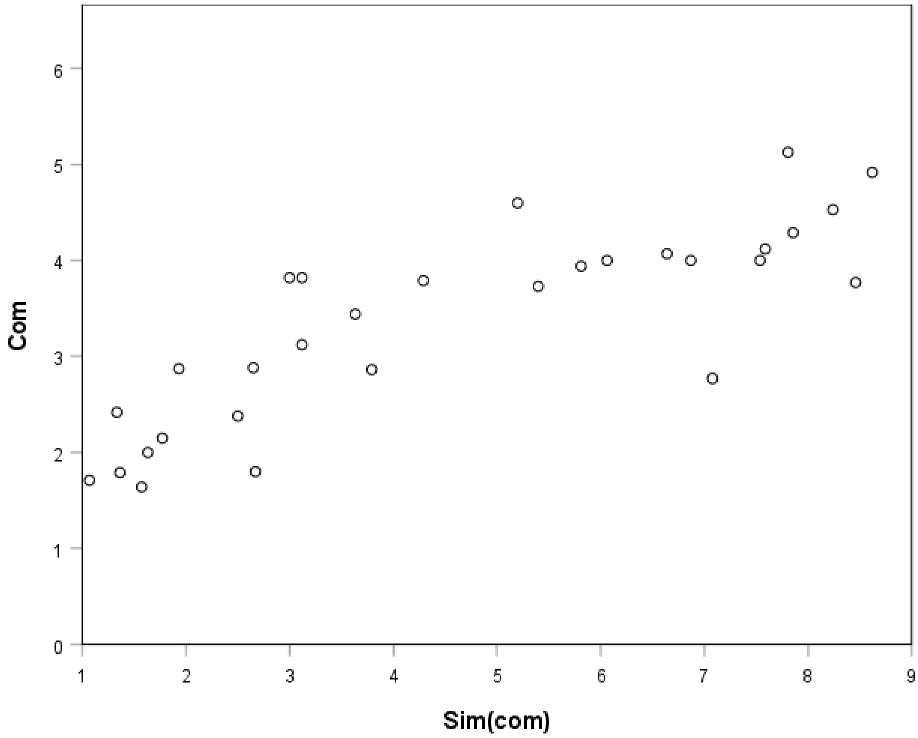


Figure 7-12 - Scatter plot of mean rated 'Sim(com)' against the mean number of listed 'Com' for 30 pairs of design concepts.

7.2.4.3 Testing hypothesis 3 – Correlational relationships between commonalities and alignable differences

S-H3 states *that concepts with many commonalities should also have many alignable differences.* That is, there should be a positive correlation between the mean number of listed commonalities and the mean number of listed alignable differences for the sample of concepts. Additionally, given that there is a positive correlation between similarity and the number of commonalities (Section 7.2.4.2), there should by extension also be a positive correlation between rated similarity and the number of alignable differences.

Commonalities and alignable differences. A Pearson's product-moment correlation was run to assess the relationship between the mean number of listed commonalities (Com) and the mean number of listed alignable differences (AD) for all 30 pairs of design concepts. Visual inspection of the scatterplot of mean Com against mean AD (Figure 7-13) indicates that there is generally a linear relation, but that there are outliers that may violate this assumption. Both Com (p = .101) and AD (p = .476) were normally distributed. There was no statistically significant correlation between the number of listed commonalities and the number of listed alignable differences, $r(28) = -.302$, p = .052 (1-tailed), with 9.1% of the variance explained. The direction of the association is also contrary to the expected positive correlation.

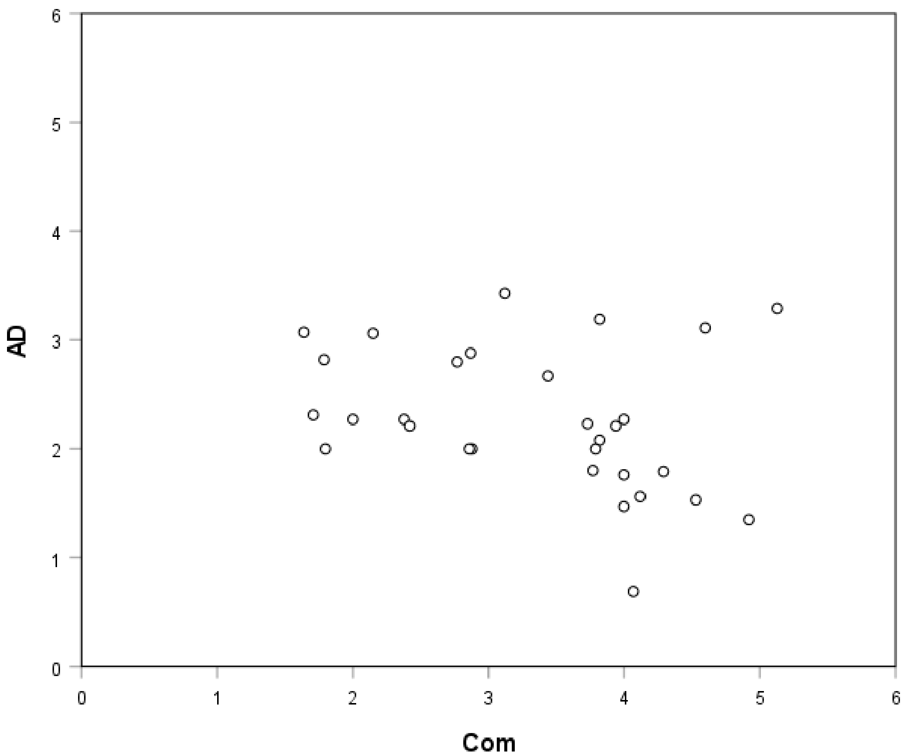


Figure 7-13 - Scatter plot of the mean number of listed 'Com' against the mean number of listed 'AD' for 30 pairs of design concepts.

Similarity and alignable differences. The Structural Alignment model predicts a positive correlation between similarity and the number of listed alignable differences by extension of the positive correlation between Com and AD. A Pearson's product-moment correlation was run to assess the relationship between mean Sim(dif) and mean AD for all 30 pairs of design concepts. The mean number of listed AD was normally distributed, (p = .476), but mean rated Sim(dif) was not (p = .012). There was a statistically significant, moderate negative correlation between rated similarity and the number of listed alignable differences $r(28) = -.471$, p = .004 (1-tailed), with 22.2% of the variance explained. This is contrary to the predictions of the Structural Alignment model but is consistent with the direction of association between Com and AD.

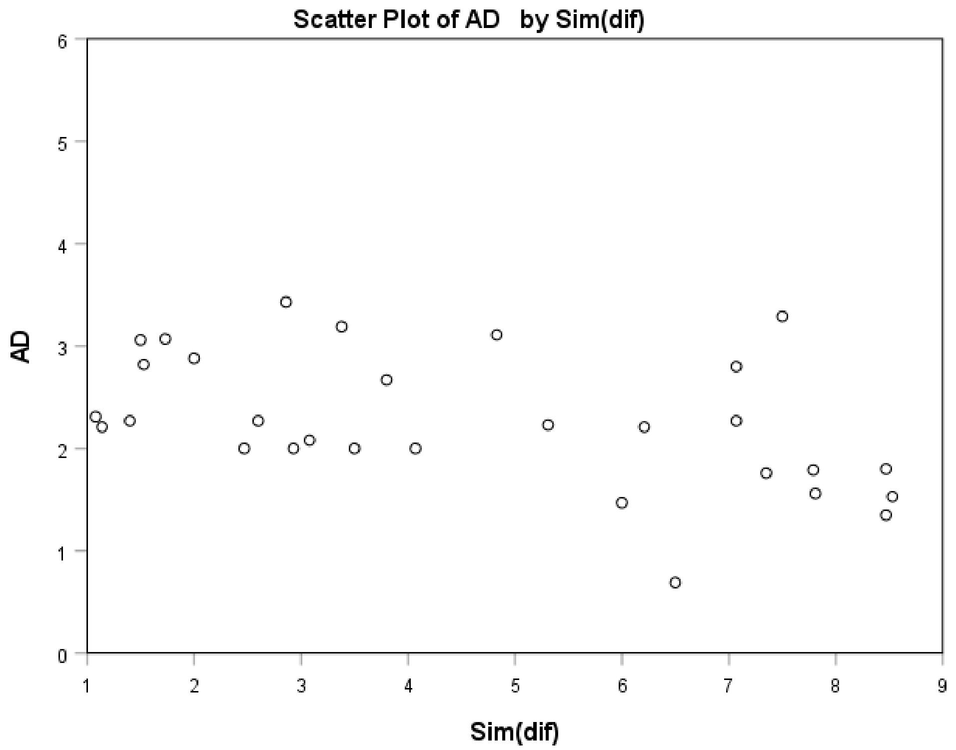


Figure 7-14 - Scatter plot of mean rated 'Sim(dif)' against the mean number of listed 'AD' for 30 pairs of design concepts.

Taken together, these findings show no support for H2. The key prediction, that concepts with many commonalities should have many alignable differences, was not supported. Nor was the related secondary prediction that rated similarity should increase with the number of alignable differences. Contrary to expectations, a moderate *negative* correlation was found between similarity and alignable differences.

7.2.4.4 Testing hypothesis 4 – Number of alignable versus nonalignable differences

S-H4 states that *alignable differences should be more numerous than nonalignable differences.* That is, participants should list more alignable differences than nonalignable differences on the whole. This prediction can be tested with individual-level data since the alignable and nonalignable differences come from the same participant. For the entire sample of responses to 30 pairs of design concepts, there were 451 responses to the listing task in which a participant listed a difference and these differences have been coded as alignable or nonalignable. It was expected that there should be statistically significantly more alignable differences listed than nonalignable differences across the entire sample.

To determine the appropriate statistical technique for testing hypothesis 3, the distribution of the difference between alignable and nonalignable differences (AD – ND) was computed. The individual-level difference scores were not normally distributed as assessed by Shapiro-Wilk's test ($p < .001$). The Shapiro-Wilk's test can be overly sensitive with large sample sizes (Field, 2009; Ghasemi and Zahediasl, 2012; Öztuna *et al.*, 2006), but visual inspection of the histogram (Figure 7-15, b) and Q-Q plot (Figure 7-15, c) reveal a tail on the data that result in a positively skewed, non-normal distribution.

Owing to the violation of normality, a Wilcoxon signed-rank test was used to determine whether there was a difference between the number of listed alignable and nonalignable differences. One of the requirements of this test is that the data are symmetrical, but the same tail on the data that make the distribution of AD-ND non-normal also violates the symmetry of the distribution of values. Thus, the Wilcoxon signed-rank test was repeated after removing the outliers shown in Figure 7-15a.

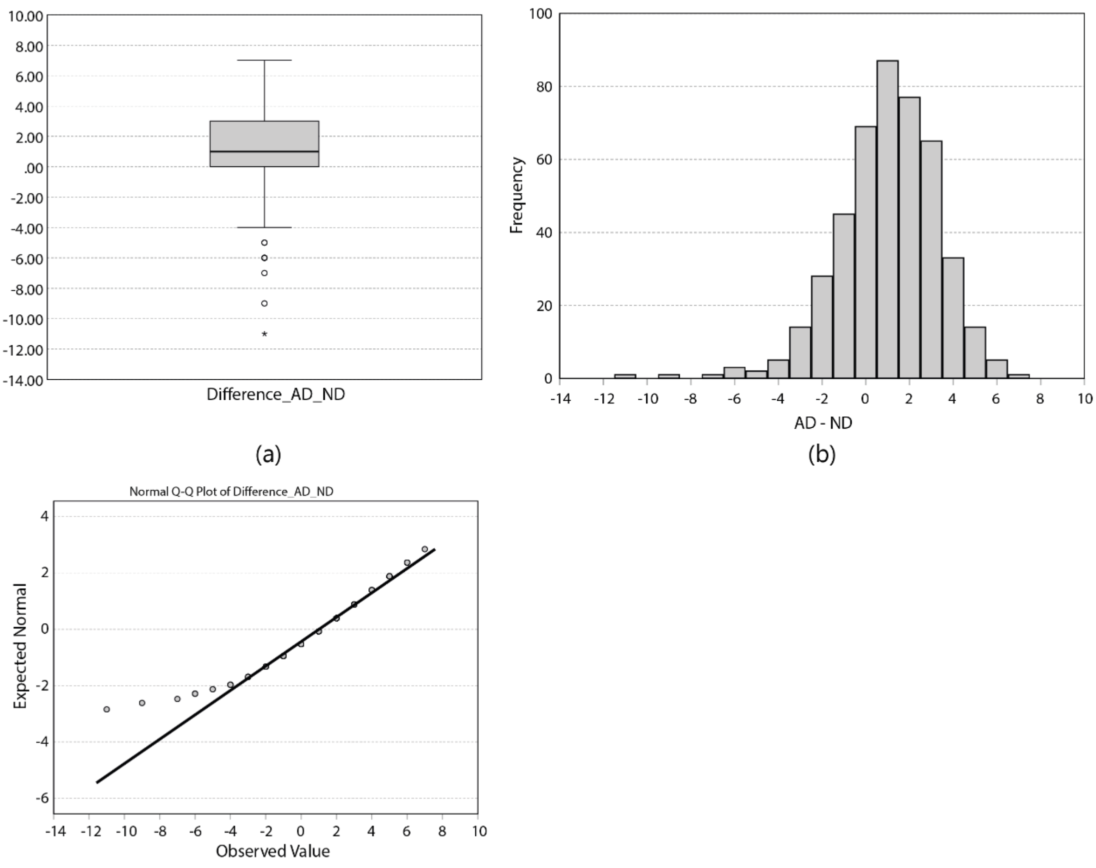


Figure 7-15 – Supporting figures for paired samples t-test for hypothesis 3. Individual-level data is used. Panels show a) boxplot of AD-ND showing outliers, b) frequency histogram of AD-ND and c) Q-Q plot for AD-ND.

Alignable differences were more numerous than nonalignable differences in 282 cases, nonalignable differences were more numerous in 100 cases, and the same number of ADs and NDs were listed in 69 cases. Participants listed a statistically significantly greater number of ADs (Mdn = 2.0) than NDs (Mdn = 1.0), W = 16,719, $p < .001$, z = -9.279. This effect is robust to removal of the outliers indicated by the boxplot in Figure 7-15 (a), n= 445, W = 13,728, $p < .001$, z = -10.296.

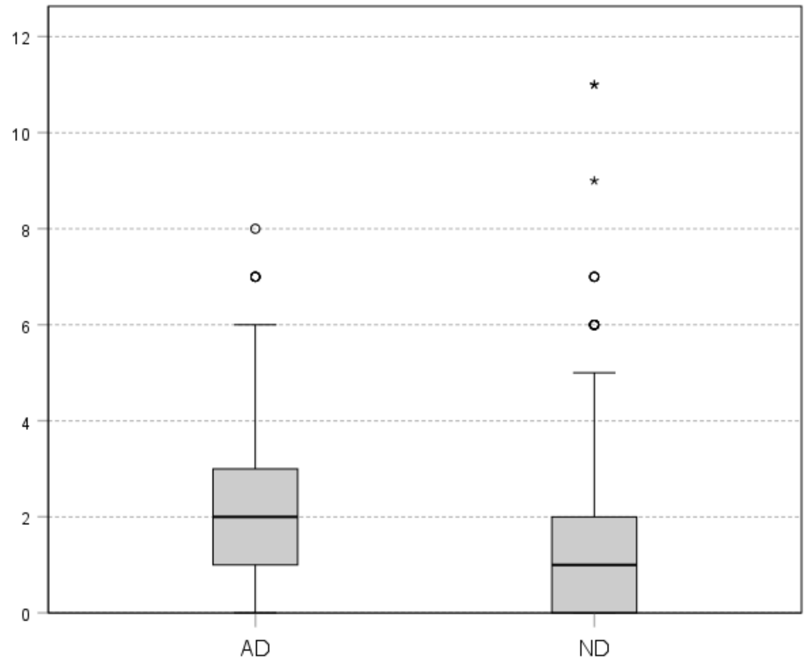


Figure 7-16 - Boxplot of the number of listed: alignable differences (AD) and nonalignable differences (ND), for all 452 responses (no outliers removed) to the listing task.

These findings support H3, that alignable differences should be more numerous than nonalignable differences. This is consistent with the structural alignment proposal that alignable differences should be more salient than nonalignable differences.

One concern with this finding stems from the fact that the base concepts for DT07 received relatively more total differences and alignable differences than those in DT03 and DT06. It is possible that the support for H4 is largely driven by the extra alignable differences listed for DT07. To address this possibility, the Wilcoxon signed-rank test was re-run for each design task. There were significantly more ADs listed than NDs in all three design tasks. This confirms the pattern of findings originally associated with H4 and shows that the relatively greater number of alignable differences cannot be attributed to a confounding influence from the inconsistencies in responses across the design tasks.

7.2.5 Discussion of Sim-P3

Five hypotheses were derived from the proposition that design concept similarity judgements happen via a comparison process and that that comparison process operates via structural alignment. The predictions concern the relationship between the outputs of similarity judgements (elicited via similarity ratings) and comparison (elicited via a commonality and difference listing task). Hypotheses H1a and H2 are common predictions of featural models of similarity judgements. Support for these would be taken as evidence that design concept similarity judgments are carried out via some kind of comparison process, but would not aid in determining the nature of that comparison process. If design concept similarity judgements occurred via a process of Structural Alignment, it was expected that all of the five predictions would be supported. Falsification of any single hypothesis would mean rejection of the SA model.

The test of H1a showed that when designers make similarity judgements of design concepts, the relative similarity of a pair of design concepts can be predicted by the number of commonalities and differences listed for the pair (explaining 77% of the variance) with commonalities contributing more to similarity than differences detract from it. Further, as the relative similarity of a pair of design concepts increases, so too does the number of commonalities that those concepts have (H2), as shown by correlational analyses.

The remainder of the predictions (H1b, H3 and H4) concerned the purported special nature of alignable differences over nonalignable differences in a Structural Alignment model of similarity. In accordance with this view, it was found that it is alignable differences, not nonalignable differences, that are significant predictors of rated similarity (with 76.1% of the variance explained) (H1b). Alignable differences were also found to be more numerous than nonalignable differences (H4). However, the prediction that concepts with many commonalities (i.e., those that are highly similar) should have many alignable differences and vice versa (H3) was not supported.

The lack of support for H3 is surprising in the context of the support for H1b and H4. Alignable differences appear to be easier to list than nonalignable differences (H4) and exert a greater influence on rated similarity (H1b), thus providing support for the expected special nature of alignable differences. So why then do concepts with many commonalities not also have many differences? Three possible explanations can be considered.

One explanation for the null result from H2 is the presence of a Type II (false negative) error caused by insufficient experimental power. Using an estimated expected effect size of 0.58 (which was the lowest correlation coefficient detected by Markman and Gentner (1996) for an equivalent analysis[16]), a two-sided correlation test had approximately a 93% chance of detecting a significant effect (calculated using G*Power 3 (Faul *et al.*, 2007)). Using this benchmark there was a low probability of the study being underpowered. Nonetheless, it is always possible that the true effect size in the population is smaller, and that the probability of a Type II error was higher. For example, the likelihood of finding a small to medium effect r=0.3 with n=30 is 49.7%. If the true effect was this small, then the lack of support for the Structural Alignment model could be attributable to insufficient statistical power.

A second explanation is that noise in the measurement of commonalities or alignable differences could mask the expected pattern of results. The correlational analyses use aggregate values of the number of commonalities and differences listed for each pair of base concepts. If participants were highly inconsistent in the number of commonalities or alignable differences that they listed, then the expected correlation may have been lost in the noise of the aggregate values. Unfortunately, since prior studies (Markman and Gentner, 1993b, 1996) did not report the standard deviations of their data it is not possible to evaluate the variation in the data gathered in Sim-P3. Two sources of noise could be the coding or the input method for responding to the commonality and difference listing task. For the coding, although the inter-rater reliability was deemed sufficient (kappa = 0.797), the remaining disagreement could contribute to issues of noisy measurement. For the input method, one methodological change between prior research and Sim-P3 was the use of typed responses. Prior studies had used audio responses (Markman and Gentner, 1993b, 1996), which were replaced with written responses in Sim-P3 to reduce the time required for data analysis. Typing the lists of commonalities and differences may have made the participants more inconsistent in the number of differences they listed or may have introduced some other unseen biases into their responses.

A third explanation is that statistical power and measurement are sound, but that some hidden phenomenon is masking the expected positive correlation between C and AD. If H2

[16] Prior research (Markman and Gentner, 1996) found correlation coefficients of r=0.58 and r=0.69, the smaller of these two is used here as a conservative estimate of expected effect size.

were to be supported, it would be expected that a positive, linear relationship would be found between Coms and AD (c.f Figure 7-13). An excess of alignable differences at high similarity or inhibition of alignable differences at low similarity could mask this relationship.

An excess of alignable differences could occur if participants found it just as easy to list alignable differences for low commonality pairs as they do for high similarity pairs owing to the use of some low-effort comparison process. For example, it may always be possible to identify 'generic' alignable differences in terms of behaviour. That is, all artefacts may be comparable in terms of weight, size and cost etc. and so it may be trivially easy to list a baseline number of alignable differences for all pairs of concepts. These responses could be filtered out by coding for formulaic or superficial responses (see: Gentner and Gunn, 2001).

An inhibition of alignable differences would mean that participants are not listing all the alignable differences that are available to them. This could occur if participants have access to a greater number of differences than they are listing but set a stopping rule for themselves, moving on to the next pair before creating a truly exhaustive list of differences, e.g., as a result of boredom or the desire to influence the pace of the experiment. Another reason for inhibition of alignable differences could be that each participant's mental alignment of design concepts results in a limited number of differences becoming available for listing. As discussed in Section 3.2.1.2, the Structural Alignment process operates to satisfy the constrain of one-to-one mapping, meaning that relations and their arguments in one design concept are matched to no more than one corresponding relation or argument in the other concept. It may be the case that for design concepts there are many possible mappings for any given pair of design concepts (meaning there are a wide variety of alignable differences that could be identified) but that any single alignment highlights a limited number of ADs, meaning that participants are not listing many ADs.

As H3 was falsified, the Structural Alignment model must be rejected as a model of design concept combination. Thus, no conclusions can be drawn about the algorithmic-level processes involved in design concept similarity judgements.

7.3 Chapter summary

This chapter has presented the results of two experiments (Similarity Experiments 2 and 3) carried out to test the proposed model of design concept similarity judgements. The results of the experiments are summarised in Table 7-11.

Similarity experiment 2 was carried out to determine the number, type, and prevalence of cognitive processes involved in design concept similarity judgements. Explanations provided for 439 similarity ratings were coded based on whether they included features of two design concepts or thematic relations between the two concepts. Results showed that almost all (96.13%) involved the statement of common or different features and no responses involved thematic relations.

The data from Similarity Experiment 3 was carried out to extend the findings of the previous experiment and evaluate the predictions of the Structural Alignment model of comparison-based similarity judgements. Four hypotheses were proposed, taken from the Structural Alignment model of similarity judgements (Markman and Gentner, 1993b, 1996). The results supported three of the four predictions of the Structural Alignment model. Two predictions common to featural models of similarity judgements were supported, providing additional evidence in support of a comparison process. Since one of the predictions was falsified, the model cannot be accepted as a model of design concept similarity judgements.

Table 7-11 - Summary of research questions, hypotheses and associated results for similarity experiments.

Ref.	Research question or hypothesis	Result
Similarity model evaluation		
Research questions, computational level		
RQ – S1	Are explanations for design concept similarity ratings indicative of (i) a comparison process and (ii) a scenario creation process?	Evidence of featural explanations (indicating comparison process) but no thematic explanations (indicating scenario creation process)
RQ – S2	What is the relative prevalence of each explanation type?	Explanations were exclusively featural
RQ – S3	What is the relationship between concept pair similarity and the type of similarity explanation?	Question not applicable owing to lack of thematic responses.
Hypotheses, algorithmic level		
H – S1a	Similarity should increase as a function of commonalities and decrease as a function of differences and commonalities should influence similarity more than differences.	**Supported** - via regression analysis (Section 7.2.4.1).
H – S1b	Alignable differences should be more important in evaluating similarity comparisons than nonalignable differences (Markman and Gentner, 1996).	**Supported** via regression analysis, but see the discussion about the role of nonalignable differences in (Section 7.2.4.1).
H – S2	Similar concepts should be associated with an increased number of commonalities and dissimilar concepts should be associated with a decreased number of commonalities.	**Supported** - via the strong positive correlation between similarity and commonalities and the strong negative correlation between similarity and differences (Section 7.2.4.2).
H – S3	There should be a numerical link between commonalities and alignable differences.	**Not supported** – there was no statistically significant correlation between Com and AD and the direction of association was negative which is contrary to expectations. There was a significant negative correlation between Sim(dif) and AD which is again contrary to expectations (Section 7.2.4.3).
H – S4	Alignable differences should be more numerous than nonalignable differences.	**Supported** - significantly more alignable differences were listed than nonalignable differences across the entire sample (Section 7.2.4.4).

8 RESULTS PT. II: DESIGN CONCEPT COMBINATION

This chapter presents the analysis and results of Combo-P3 to satisfy Obj4 – to *evaluate the hypothesised Dual-Process model of design concept combination.* Satisfaction of this objective involved addressing three research questions about the inputs and outputs of combination, and testing two hypotheses that make predictions about the relationship between similarity and combination type.

Combination experiment 1 (Combo-P1) was conducted as a pilot study to test the experimental procedure for the combination task, and so no results from this experiment are reported in this chapter. As shown in the research map (Figure 4-6), there was no combination experiment in Phase 2 of the research, so there is no Combo-P2.

The research questions associated with the combination model were addressed using the measures of combination type and trial difficulty from Combo-P3 and base-concept similarity from Sim-P3. The answers to the research questions meant that the planned hypothesis tests could not be carried out. This is discussed in Section 8.1.5.2. The results are discussed in Section 9.2 to evaluate the proposed model of design concept combination.

8.1 Combo-P3

The four research questions associated with the proposed Dual-Process model of design concept combination are listed again in (Table 8-1). Combo-P3 involved 30 designers combining 30 pairs of design concepts to create new design concepts that addressed the same brief. The designers provided difficulty ratings after responding to each stimulus. The same stimuli were used in all experiments in Phase 3 of the research, meaning that the design concepts that were combined in Combo-P3 were the same design concepts that were used in the similarity rating and commonality and difference listing task in Sim-P3.

Table 8-1 – Research questions associated with the hypothetical model of design concept combination and the section in which they are addressed

Research question
RQ – C1: What types of combinations do designers produce?
RQ – C2: What is the prevalence of each combination type?
RQ – C3: What is the relationship between concept pair similarity and the type of combined concept?
RQ – C4: What is the relationship between concept pair combination difficulty and the type of combined concept?

8.1.1 Analytic considerations and definition of variables

8.1.1.1 Selection of statistical tests

The analyses conducted to answer RQ-C3 and RQ-C4 concern the relationship between a continuous variable (similarity ratings) and proportional data (the relative proportion of combination types). The similarity ratings and combinations were produced by independent samples, thus necessitating the use of statistical tests for independent samples. To examine the differences in proportions across multiple groups, Chi-square tests of independence were used (Laerd Statistics, 2017). To examine the relationship between similarity and proportion types, correlational analyses were used, with the caveat that proportional data can give rise to spurious correlations, i.e., a positive association between A and B cannot be disentangled from a negative association between A and C.

8.1.1.2 Definition of variables

Table 8-2 lists the variables that were used to code the combination types. The outputs of the design concept combination task were coded using the variables in the ‘code’ column. The codes are grouped according to whether the designer successfully ideated or not (i.e., produced a design concept) and whether ideation involved combination or not (the 'response type' column. These groupings are classes of code that were defined *post hoc.*

Table 8-2 – Definitions of variables used in the analysis of Combo-P3

	Response type	Code
Ideation	C – Combination	F – Feature mapping
		R – Relational
		A – Ambiguous
	Nc – Non-combinational	Nc – Non-combinational
No ideation	O – Other	I – Insufficient information to interpret sketch
		U – Unclear how concept addresses brief
		NR – No response (could not generate)

The use of proportional data raises the question of which output types should be analysed together. For example, the research questions are concerned with the relationship between similarity and *combination type,* i.e., the proportion of F, R, and A codes. However, almost 20% of the responses involved successful instances of ideation but contained no evidence of combination (Nc codes). Although it is the F, R and A codes that are theoretically relevant, ignoring the Nc codes might exaggerate or underrepresent patterns in the data that may look different in the context of the full sample of responses. To address this, some analyses were conducted multiple times using one of three different groups of variables.

(i) Response types (C, Nc, O). These variables describe whether the designer ideation and combined, ideated but did not combine or failed to ideate.

(ii) Ideation types (F, R, A, Nc). These variables describe all successful instances of ideation. This group of variables enables analyses of combination types whilst avoiding the risk that excluding the Nc responses might skew the results.

(iii) Combination types (F, R, A). These variables describe only ideation in which combination occurred. These are the variables that are theoretically relevant for the proposed Dual-Process model of design concept combination.

Two types of nomenclature are used. Single letters (C, Nc, O, F, R or A) represent the count of the respective variable. Letters with subscripts represent the proportion of the class of variables described in the subscript. For example, $C_{(response)}$ is the number of combinations as a proportion of response types (C, Nc, O), $F_{(combo)}$ is the number of F codes as a proportion of all combination type codes (F, R and A), $F_{(ideation)}$ is the number of F codes as a proportion of all ideation types (F, R, A, Nc).

8.1.2 Descriptive statistics and example responses

The descriptive statistics provide summary data for the subsequent analyses and directly address RQ-C2 by showing the prevalence of the featural and relational combination types.

8.1.2.1 Example responses

Examples of featural, relational and ambiguous responses are provided below. Figure 8-1 shows a pair of base-concepts Sc_25. The brief for this design task was to “generate concepts for products that may facilitate water transportation in developing nations”. Concept Sc_25a comprises a bus with a roof-mounted water container that resembles an air mattress. Concept Sc_25b comprises a “heavy lift drone” with four sets of propellors that carries bottles (presumably containing water) for “mass delivery”.

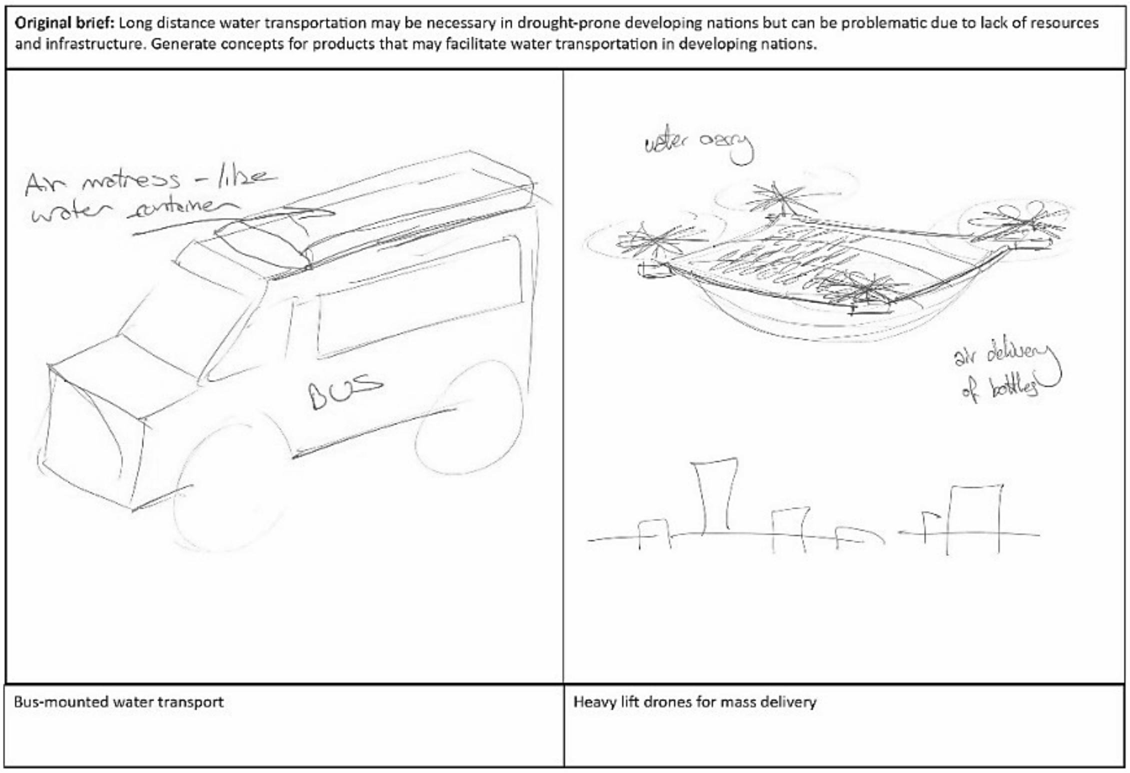


Figure 8-1 – Base concepts Sc_25a and b, used in both Sim-P2 and Combo-P3

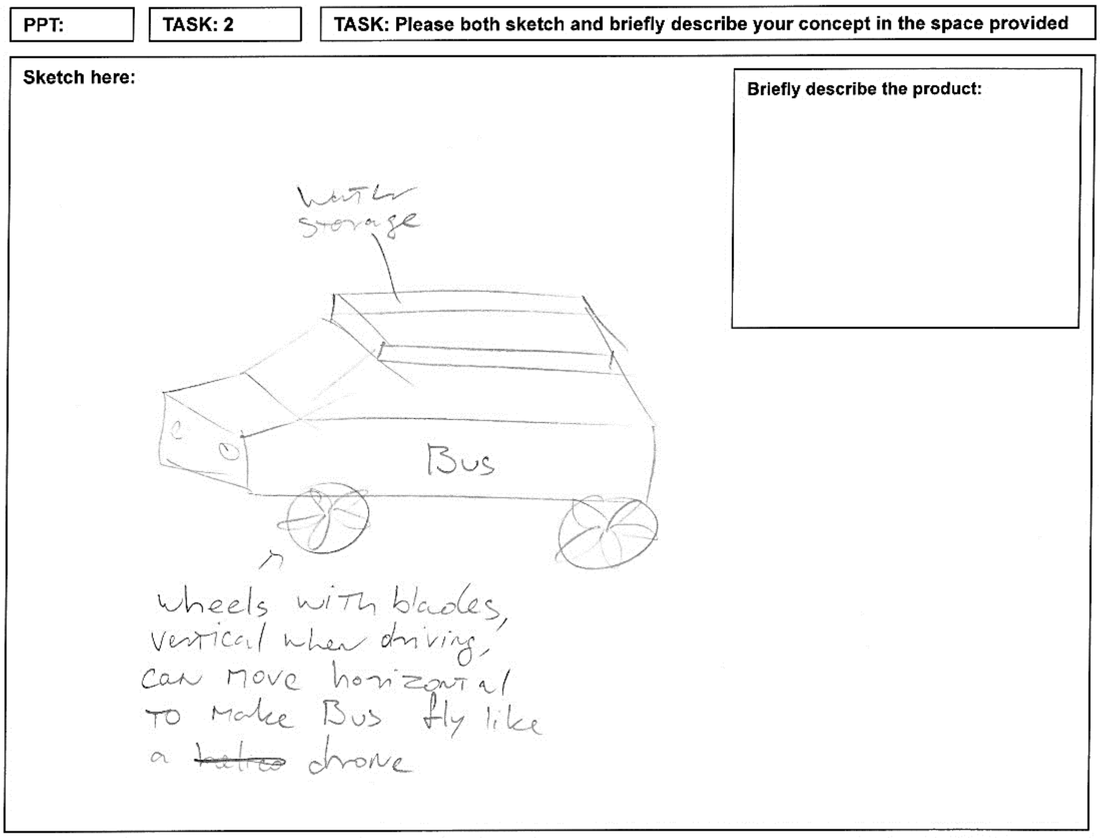


Figure 8-2 – An example of a featural combination in response to Sc_25. The description reads “wheels with blades, vertical when driving, can move horizontal to make bus fly like a drone”.

Featural combination (Figure 8-2). A bus with roof-mounted water storage and “wheels with blades” that enable the bus to “move horizontal to make the bus fly like a drone”. This appears to be concept SC_25a with features from SC_25b, such as the propellor blades, ability to fly and the likeness of being a drone.

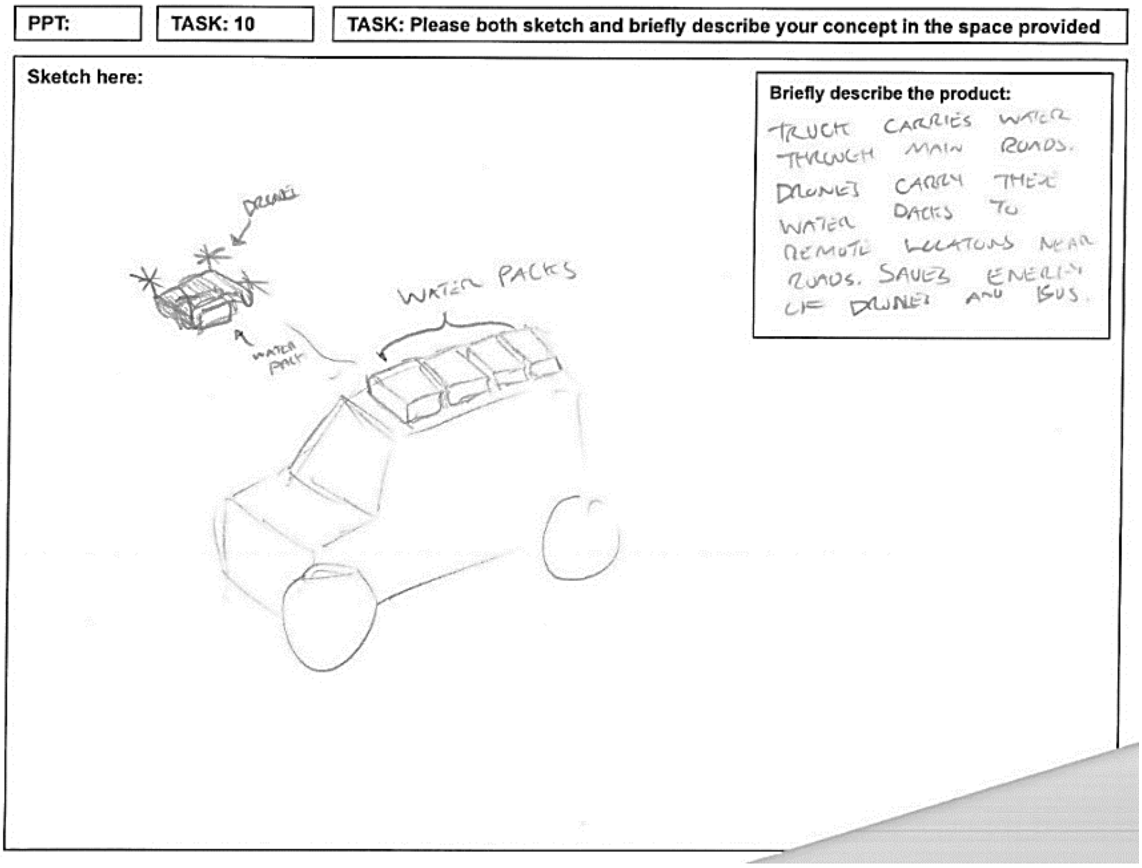


Figure 8-3 – An example of a relational combination in response to SC_25. The description says “truck carries water through main roads. Drones carry the water packs to remote locations near roads. Saves energy of drones and bus”.

Relational combination (Figure 8-3). A vehicle with roof-mounted water packs and a drone carrying a water pack at a remote location from the vehicle. The caption reads: “Truck carries water through main roads. Drones carry the water packs to remote locations near roads. Saves energy of drones and bus.” Here, the bus from SC_25a and the drone from SC_25b have been placed into a complementary system in which the bus carries the water tanks via roads and the drones take the water packs via the air to off-road locations.

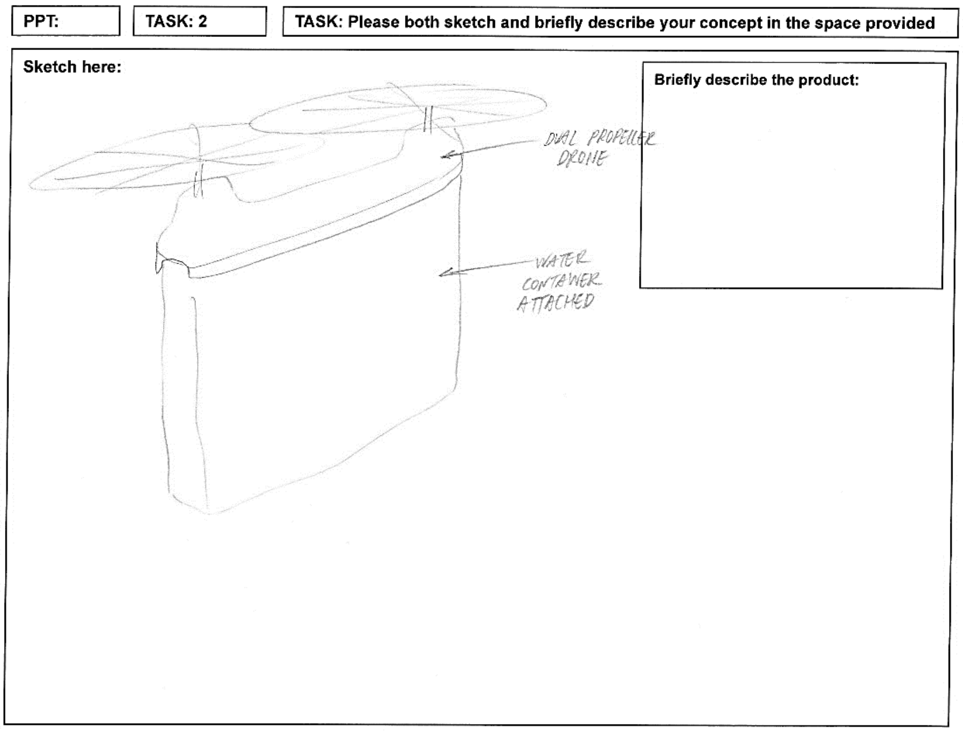


Figure 8-4 – An example of an ambiguous combination in response to Sc_25.

Ambiguous combination (Figure 8-4). The sketch comprises a “dual propellor drone” with a “water container attached”. This can be interpreted in two ways. In one interpretation the participant may have taken the water container from the bus in Sc_25a and related it through an external ‘attachment’ relation to the drone in Sc_25b, resulting in a drone with a water container attached. In a second interpretation, the designer may have transferred features of the water container in SC_25a to the bottles that are carried by the drone in SC_25b, resulting in a drone with a water container attached. In either interpretation, the drone has undergone some conceptual change to reduce the number of propellor units from 4 to 2 and to align with the geometry of the water container. Ambiguous combinations appear to occur when there are multiple, distinct entities in one of the base concepts one of those entities is ‘swapped’ for the other base concept.

8.1.2.2 Responses and exclusions

30 participants participated in Combination Experiment 2. One participant was excluded as they selectively skipped 11 of 30 trials, 9 of which were from a single design task. 22 participants completed all 30 trials, but 8 participants exercised their right to leave the experiment before completing the experiment. This reduced the total number of responses but because the stimulus presentation order was randomised, the loss was spread across the stimuli. Of a total of 870 possible responses (29 participants x 30 stimuli), participants

attempted 782 (excluding the 88 lost through early finishes). The mean number of responses to each stimulus pair was 26.07 (Table A8-1).

Participants provided 778 difficulty ratings from a possible maximum of 782. Four difficulty ratings were lost due to data entry errors on the part of the participants. No responses to the difficulty rating were considered outliers and none were removed.

8.1.2.3 Coding summary

Table 8-3 shows the number and percentage of the codes based on the combination coding scheme (Section 5.3.2.2) and Figure 8-5 illustrate the raw data for the combination types in a square area chart. The participants are listed in rows, the stimuli are listed in columns, and the type of response is indicated by the colour of the squares at the intersection of the row and column. The 'other' codes are collapsed into a single code to improve readability.

Ideation (successful creation of a design concept) occurred in 90.1% of the 782 attempted trials. Combination took place in 70.5% of the total attempted trials. 19.6% of the total attempts involved some kind of non-combinational ideation such as the generation of a new concept from scratch or the modification of one of the stimulus concepts.

Of those instances in which participants successfully ideated (n=704), 42.46% were featural combinations, 18.79% were relational combinations, 9.21% were ambiguous and 19.57% were non-combinational.

Table 8-3 – Coding summary for Combo-P3

	Ideation indicator	Response type	Code	n	*% of attempted*
All responses 782 *(100%)*	Ideation 704 *(90.1%)*	C - Combination 551 *(70.5%)*	F – Feature mapping	332	*42.5*
			R – Relational	147	*18.8*
			A - Ambiguous	72	*9.2*
		NC - Non-combinational 153 *(19.6%)*	Nc – Non-combinational	153	*19.6*
	No ideation 78 *(9.9%)*	O - Other 78 *(9.9%)*	I – Insufficient information to interpret sketch	40	*5.1*
			U – Unclear how concept addresses brief	5	*4.2*
			NR – No response (could not generate)	33	*0.6*

8.1.2.4 Summary statistics for similarity and combination type

The analyses of the relationship between similarity and combination type (RQ-C3) use the median similarity ratings from Sim-P3 and the combination types from Combo-P3. Figure 8-5 illustrates the relevant data, with numerical values listed in Table A8-1 in the appendix. The top of the figure shows boxplots of Sim(all) for each stimulus pair ordered by descending median similarity. Mean similarity is overlaid in a line chart. The bottom of the figure shows the type of response given by each participant in Combo-P3 in response to that stimulus in the combination task in a square area chart but with the columns re-ordered by rank median similarity to match the boxplots. The figure is also split into three levels of similarity (low (1 - 3.66), medium (3.67 to 6.33) and high (6.34 - 9)) according to median rated similarity to show the distribution of the responses across the similarity scale. This is indicated by the horizontal brackets above and below the squares.

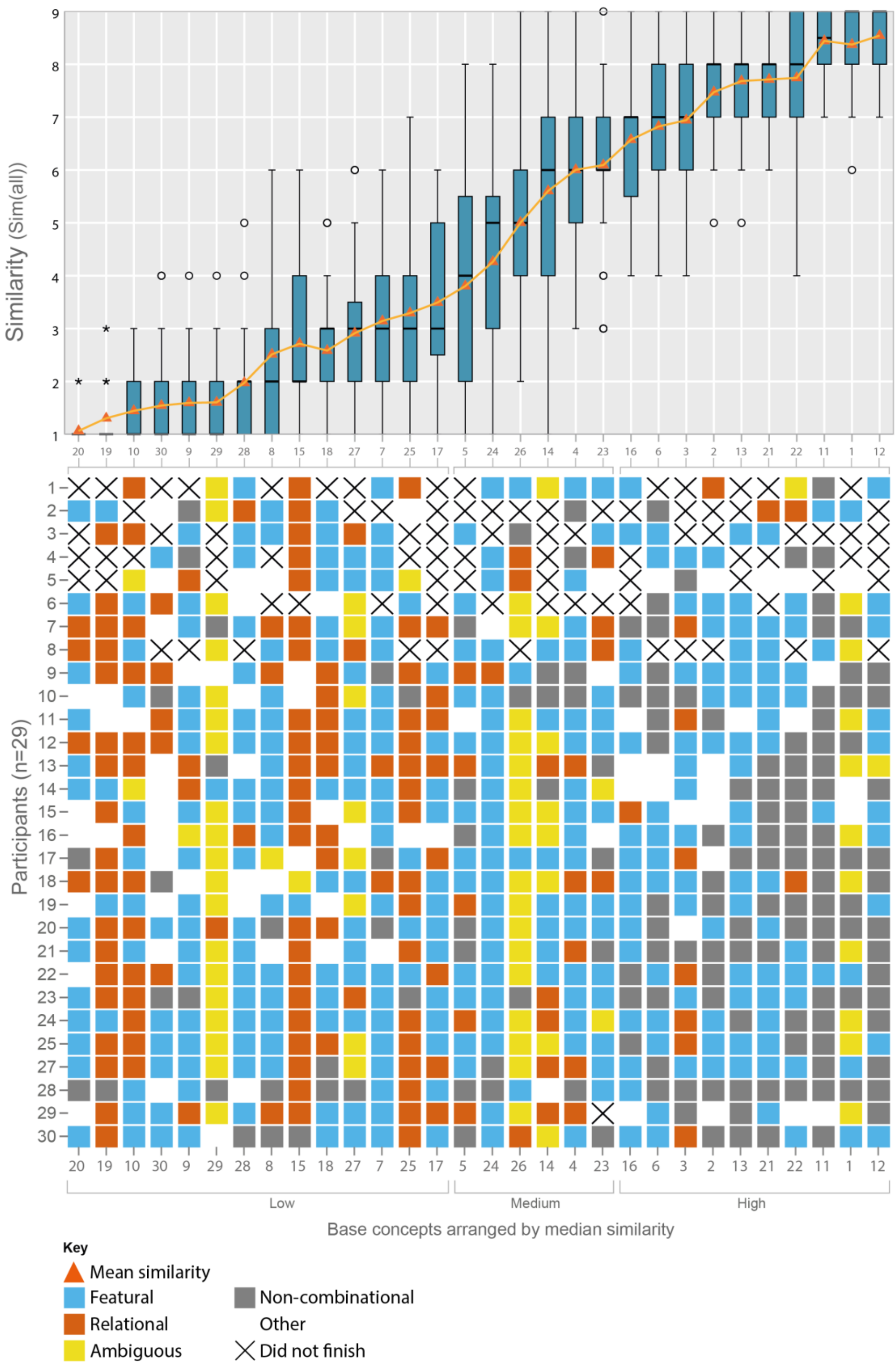


Figure 8-5 – Data summary for the similarity ratings and types of response to the combination task, showing boxplots of similarity ratings, mean similarity and a square area chart of combination response type.

8.1.3 Pre-analysis data checking

Before answering the research questions, analyses were conducted to determine whether:

- There was a confound between the design task and the proportion of response types, the proportion of combination types, or trial difficulty.
- Trial difficulty was confounded with similarity, response type or combination type.

8.1.3.1 Testing design-task homogeneity

Chi-square tests of independence were conducted to determine whether there was an association between the three design tasks and the proportion of response types and combination types. The null hypothesis for each test is that the proportion of each outcome measure is not associated with the design task. A rejection of the null would indicate that there *is* an association between the proportion of response type and the design tasks.

Design tasks and proportion of response type (Figure 8-6). There was no statistically significant association between design task and response type $\chi^2(4) = 6.073$, $p = 0.194$ and the association was small (Cohen, 1988), Cramer's V = .062. Thus, there is no confounding association between design task and response type.

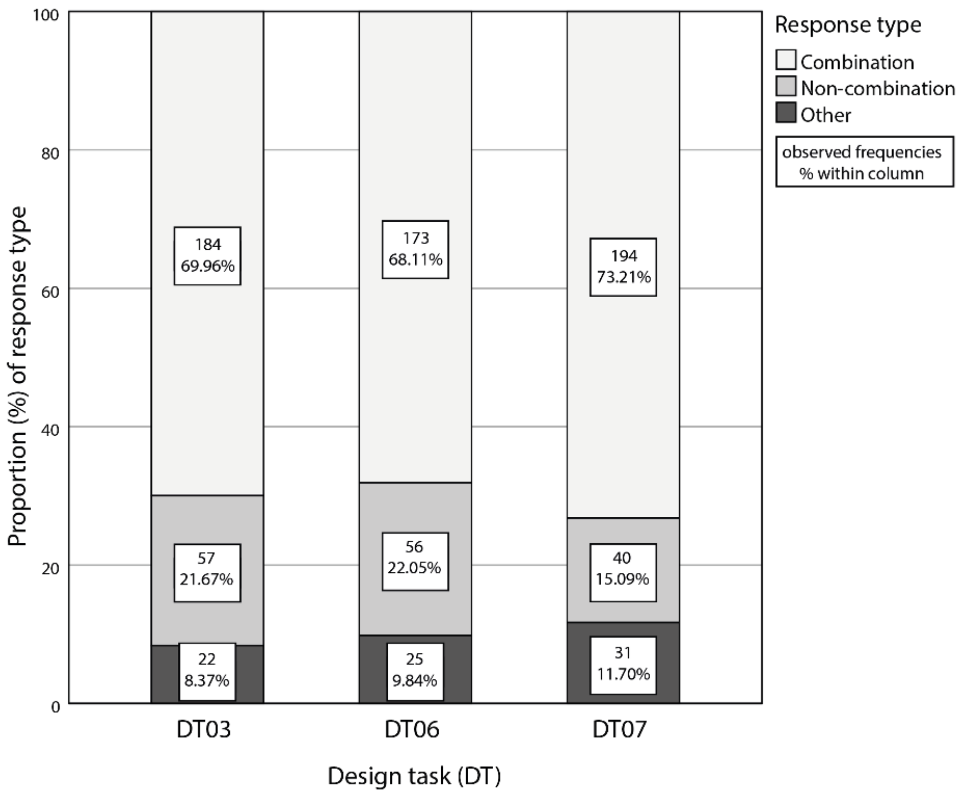


Figure 8-6 – Proportions (%) of response types per design task.

Design tasks and proportion of ideation types (Figure 8-7). There was a statistically significant association between design task and the proportion of combination types when including non-combinational responses $\chi^2(6) = 54.920$, $p < .001$. This association was small to moderate (Cohen, 1988) as indicated by Cramer's V = .197. Adjusted residuals of greater than 3 are considered to deviate significantly from independence (Agresti, 2002; Agresti and Franklin, 2014) and are flagged (*) in Figure 8-7. The magnitude of the adjusted residuals, alongside visual inspection of Figure 8-7, shows that the effect is driven by (i) an increased proportion of relational combinations in DT06 and (ii) a relative decrease in the proportion of ambiguous combinations in DT06 and an increase in DT07.

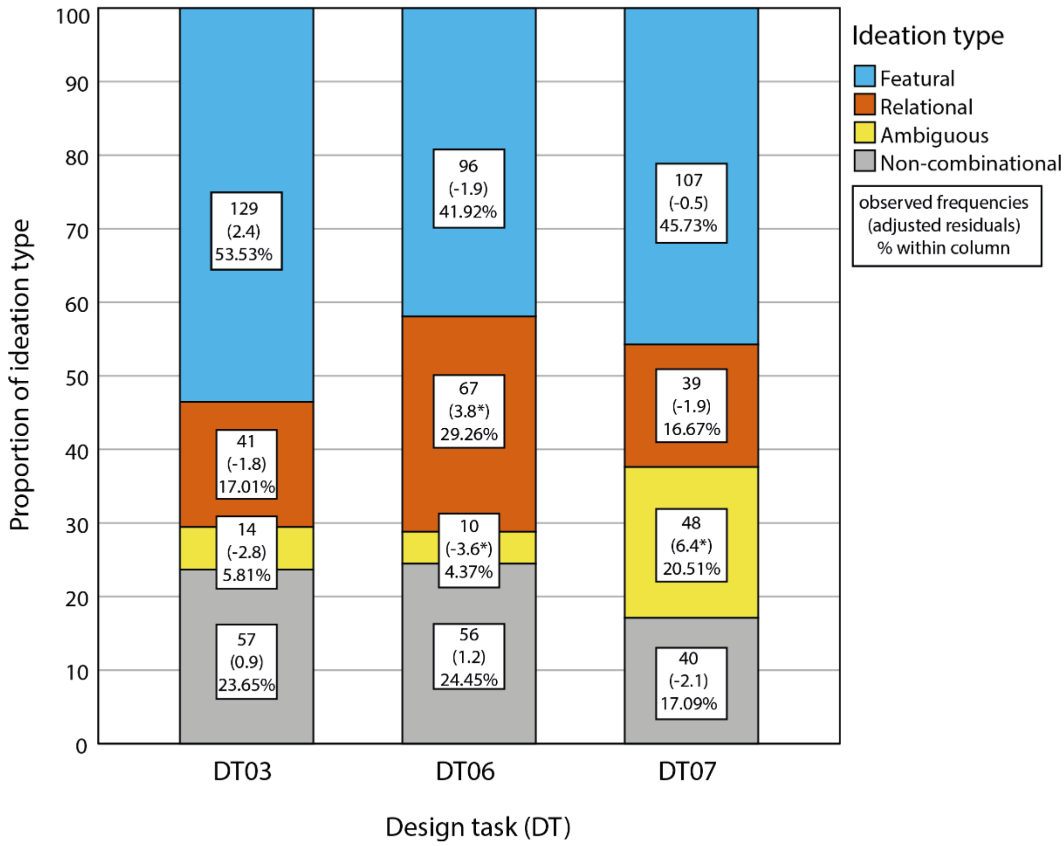


Figure 8-7 – Proportions (%) of combination type per design task when including non-combinational responses.

Design tasks and the proportion of combination types. As the previous association between design task and ideation type was not attributable to the non-combinational responses, it was expected that the third chi-square test would reveal a significant association between design task and combination type *when looking only at instances of combination*. As expected, the association was significant $\chi^2(6) = 49.808$, $p < .001$. This association was small to moderate (Cohen, 1988) as indicated by Cramer's V = .213. Adjusted residuals of greater than 3 are flagged (*) in Figure 8-8. Consistent with the previous test, the effect is driven by (i) an increased proportion of relational combinations in DT06 and (ii) a relative decrease in the proportion of ambiguous combinations in DT06 and an increase in DT07. Additionally, there was also a relative increase in featural responses for DT03.

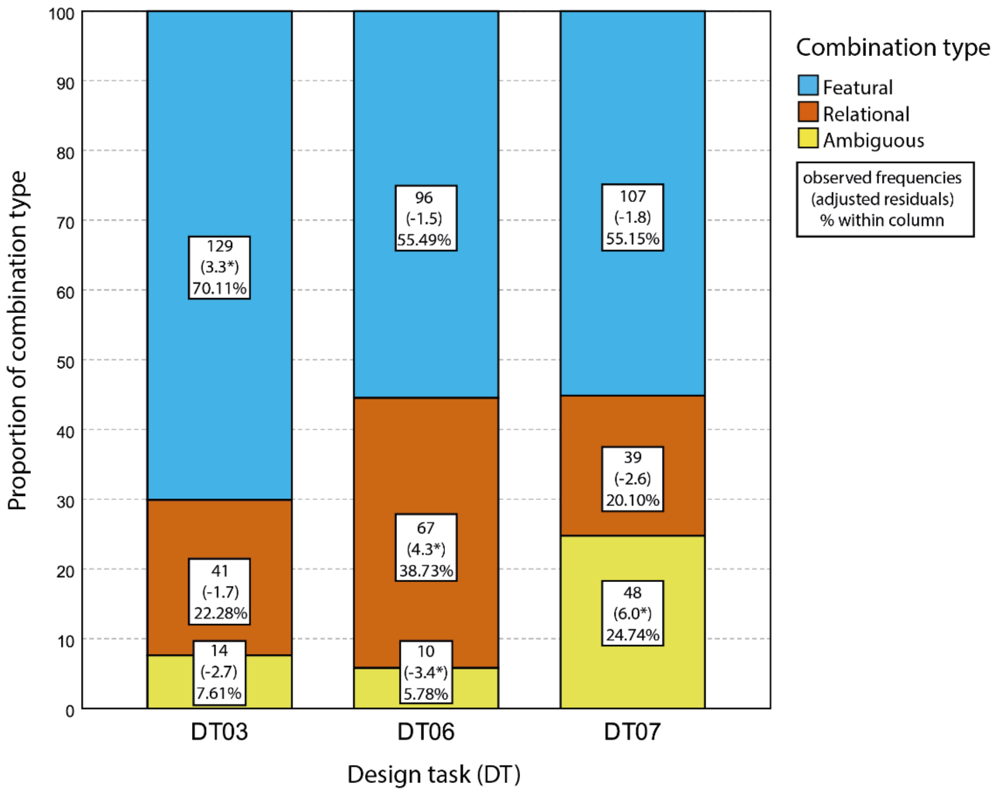


Figure 8-8 - Proportions (%) of combination type per design task.

Difficulty across design tasks. To determine whether the design tasks were equally difficult to respond to, individual-level data are used to test for differences in rated difficulty. Table 8-4 shows the median value of rated difficulty for each design task and Figure 8-9 shows (a) the boxplots and (b) the histogram of the difficulty ratings for each design task.

Table 8-4 – Median and number of responses for rated difficulty across three design tasks.

	DT03 (Pairs 1-10)		**DT06 (Pairs 11-20)**		**DT07 (Pairs 21-30)**		**Total**	
Measure	Mdn	n	Mdn	n	Mdn	n	Mdn	n
Rated difficulty	2	262	3	252	3	264	3	778

Note: Mdn = Median value, n = number of responses, DT = Design task.

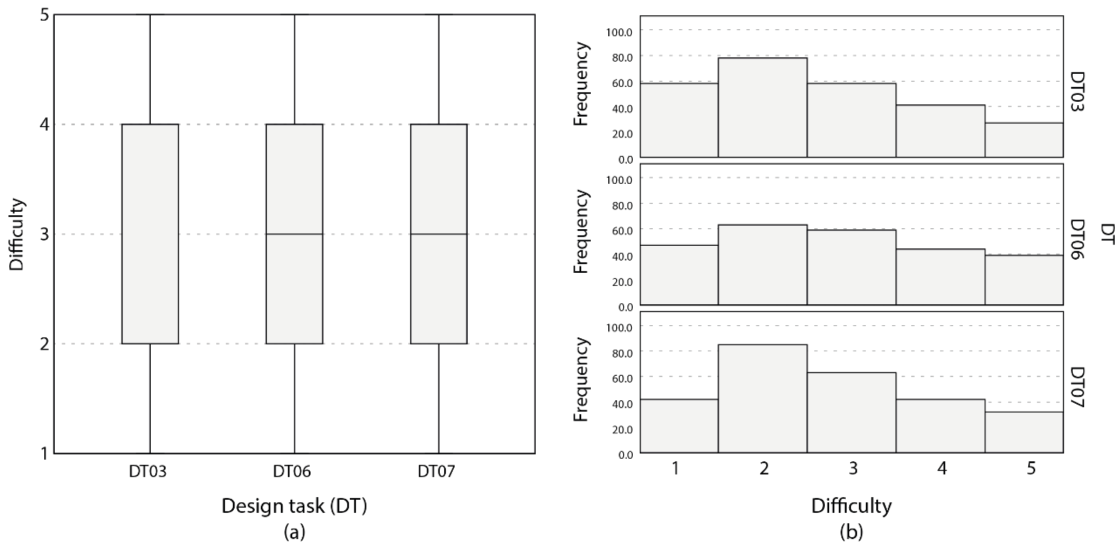


Figure 8-9 – Combination difficulty across three design tasks, showing (a) boxplots and (b) histograms.

None of the difficulty ratings were normally distributed as assessed by Shapiro-Wilk's test, ($p < .001$) and as is supported visually in Figure 8-9b. The non-parametric Kruskal-Wallis H test was used to compare the median value of responses across design tasks. The median ratings for difficulty across the design tasks were not statistically significant different χ^2 (2) = 4.334, p = .115.

8.1.3.2 Testing for difficulty confounds

Three analyses were conducted to determine whether any of the outcome measures (similarity, combination type and response type) were confounded with trial difficulty. A linear association between difficulty and any of these variables would indicate the presence of a confound.

Difficulty and similarity. Visual inspection of the scatterplot between mean similarity and mean difficulty (Figure 8-10) for each of the 30 stimulus pairs suggests that the variables are not linearly related, thereby indicating that similarity is not confounded with difficulty. The relationship between the variables was estimated using the 'curve estimation' tool in SPSS to apply a linear and quadratic model to the data. This revealed a significant linear regression model R^2=0.186 F(2, 28) = 6.398*, p = 0.017* and a significant quadratic regression model R^2=0.365, F(2, 27) = 7.774*, p = 0.002* with a u-shaped curve, both represented by the equations in Figure 8-10. The quadratic model provides a better fit for the data, supporting the observation that difficulty is not linearly related to similarity and thus is not a confounding variable.

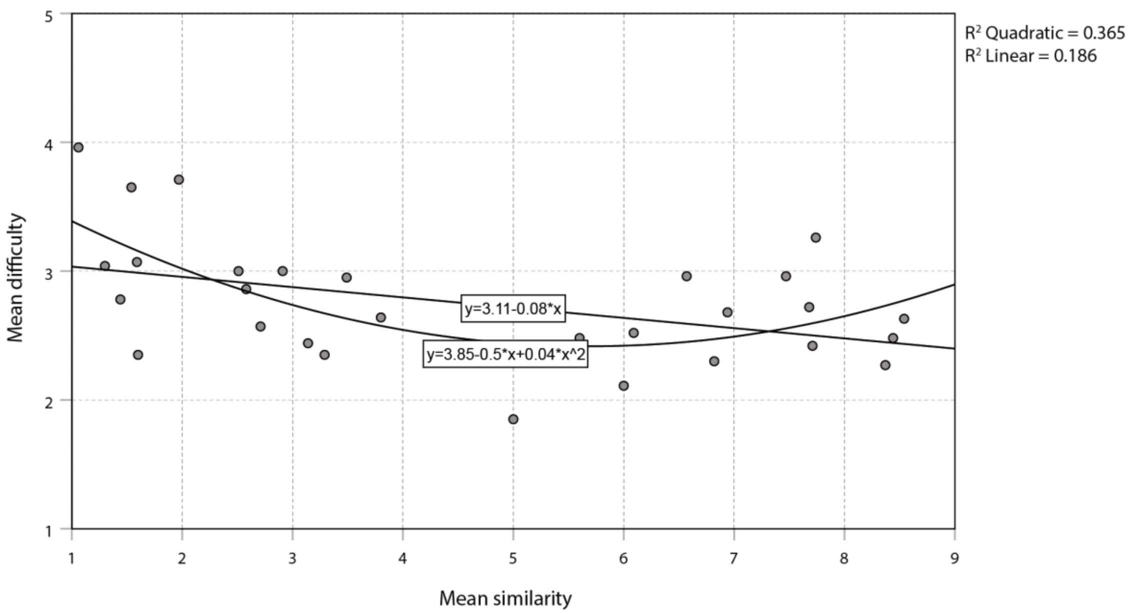


Figure 8-10 – Scatterplot showing mean rated similarity and mean rated difficulty for the 30 stimuli.

Difficulty and proportion of response type. Figure 8-11 shows the scatterplots for mean difficulty and the proportion of the three response-type codes. Visual inspection of the scatterplots shows an approximately linear, monotonic relationship between difficulty and $O_{(response)}$, but not for the other response type variables. A Spearman's rank-order correlation was run to assess the relationship between mean rated difficulty and the proportion of each response type.

- There was no statistically significant correlation between mean rated difficulty and the proportion of Combination responses $r_s(28) = -.108$, $p = .572$
- There was no statistically significant correlation between mean rated difficulty and the proportion of Non-combination responses $r_s(28) = -.277$, $p = .138$
- In contrast, there was a statistically significant, medium positive correlation between mean rated difficulty and the proportion of 'other' responses $r_s(28) = .531$, $p = .003$

These findings show that the tendency to produce combined or non-combinational design concepts is independent of trial difficulty, but as difficulty increases, participants are more likely to fail to ideate (i.e., produce an 'other' response).

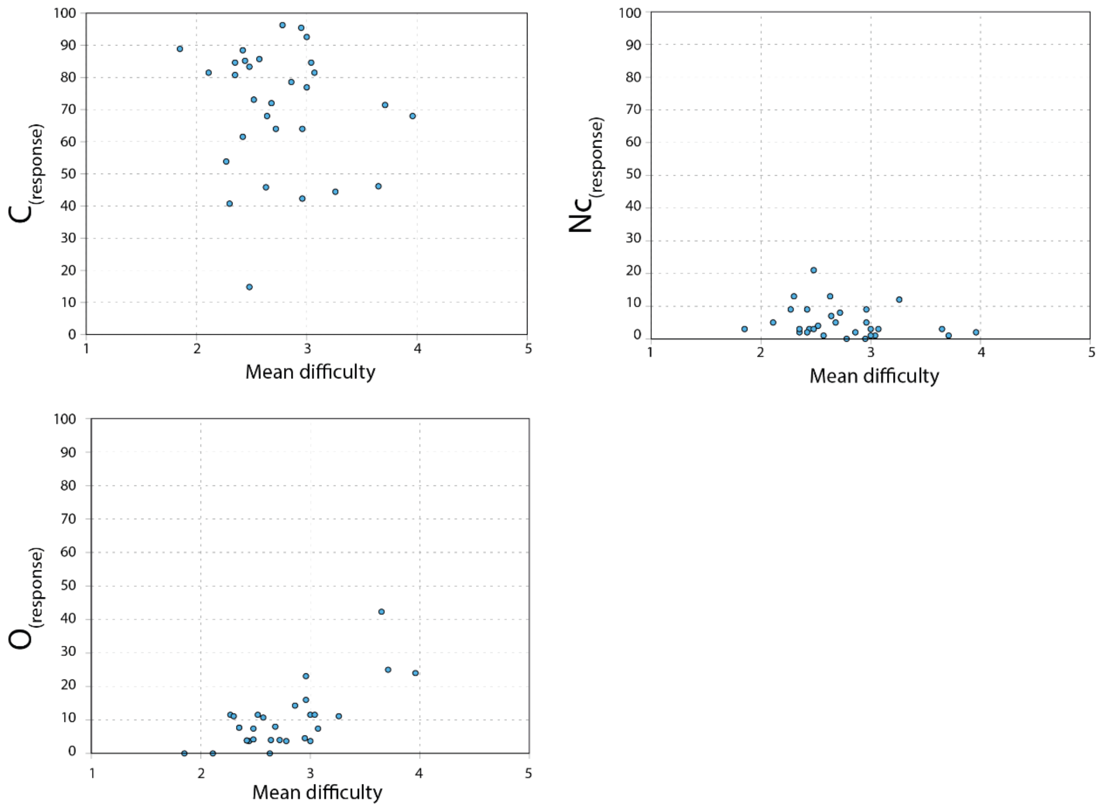


Figure 8-11 - Scatterplots showing mean difficulty plotted against the proportion of response type: (a) $C_{(response)}$, (b) $Nc_{(response)}$ and (c) $O_{(response)}$ for each of the 30 stimuli.

Difficulty and proportion of combination type. Figure 8-12 shows the scatterplots for mean difficulty and the proportion of each ideation type. Visual inspection of the scatterplots shows that there are no linear or monotonic relationships in any of the panels. A Spearman's rank-order correlation was run to assess the relationships anyway.

- There was no statistically significant correlation between mean rated difficulty and the proportion of featural responses $r(28) = .066$, $p = .731$.
- There was no statistically significant correlation between mean rated difficulty and the proportion of relational responses $r(28) = .310$, $p = .095$.
- There was no statistically significant correlation between mean rated difficulty and the proportion of ambiguous responses $r(28) = -.244$, $p = .194$.

These findings show that, for each of the 30 stimulus pairs, the proportion of featural, relational or ambiguous combinations was independent of trial difficulty. The same pattern of findings holds if non-combinational responses are included in the analyses.

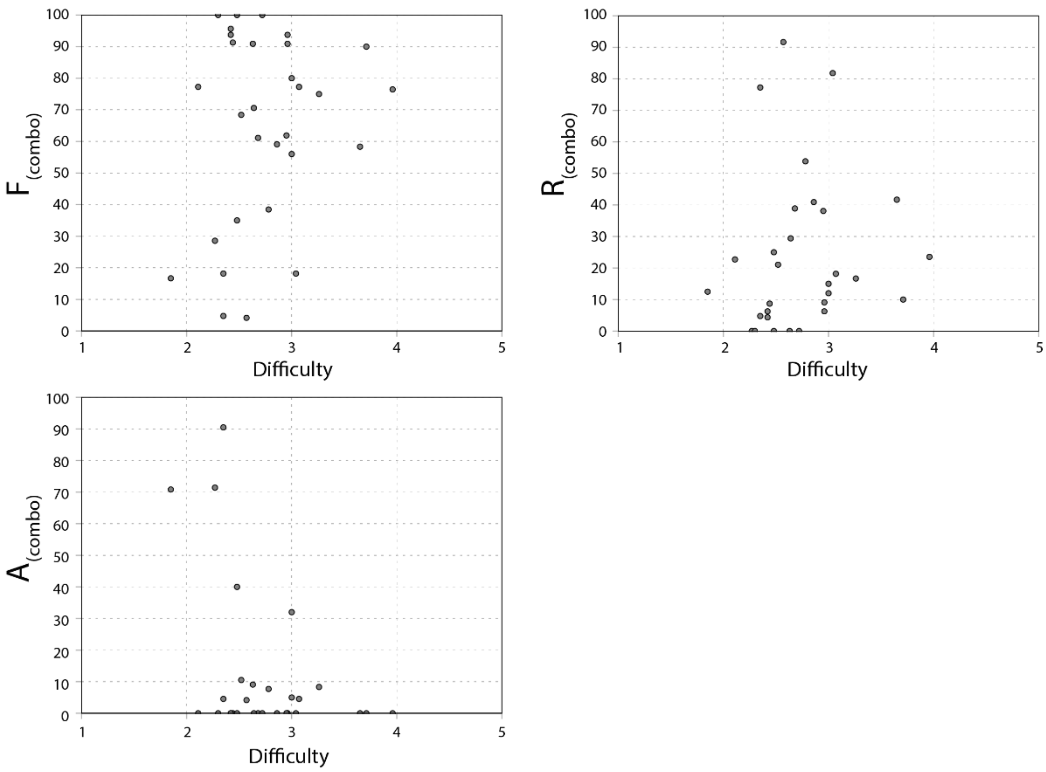


Figure 8-12 - Scatterplots showing mean difficulty plotted against the proportion of combination type: (a) $F_{(combo)}$, (b) $R_{(combo)}$ and (c) $A_{(combo)}$ for each of the 30 stimuli.

8.1.3.3 Summary

Pre-analysis data checking has been conducted to determine (i) whether responses were homogeneous across the three design tasks, and (ii) whether any of the outcome measures were confounded with trial difficulty.

With regards to the homogeneity of outcomes across design tasks, it had already been established in Section 7.2.3.1 that there were significant differences in median similarity across the three design tasks, characterised by relatively lower rated similarity in DT07 compared to DT03 and DT06. Three additional tests were conducted on the outcomes of Combo-P3.

1. **Design task and proportion of response type**. There was no significant association between the proportion of response types (C, Nc, O) and the design task.
2. **Design task and proportion of combination type**. There was a significant difference between the proportion of combination types (F, R, and A) across the three design tasks. This was attributed to: a relative increase in the proportion of featural combinations in DT03, an increase in relational combinations in DT06 and a relative decrease in the proportion of ambiguous combinations in DT06 alongside an increase in DT07.
3. **Design task and difficulty**. There was no significant difference in median difficulty across design tasks.

The association between design tasks and combination type suggests that combination type may be confounded with design task. That is, the proportion of each combination type may be partially explained by the characteristics of the three design tasks. As this does not interfere with the requirement that the stimuli span a range of relative similarity ratings 4.2.4, the decision was made to move on to the full data analysis. It does, however, have implications for construct validity. Any inferences drawn about the relationship between similarity and combination type cannot be said to apply to sets of design concepts from all design tasks. This is discussed in Section 9.2.1.3 with respect to the generalisability of the inferences about combination cognitive processing

Analyses were also conducted to check for a potential confounding effect of trial difficulty. The only statistically significant correlation was between trial difficulty and the proportion of 'Other' responses, $O_{(response).}$ As will be seen in Section 8.1.4, $O_{(response)}$ was not statistically significantly correlated with any other outcome measures and so was not a confounding variable. Further, the absence of any linear relationships between trial difficulty and any other variable shows that difficulty does not pose a threat as a confounding variable.

8.1.4 Results: combination, similarity and difficulty

A series of analyses were undertaken to address RQ-C3 and RQ-C4, which concerned the relationships between the outputs from the combination task and similarity and difficulty, respectively. All tests of normality were assessed by Shapiro-Wilk's test unless otherwise stated.

8.1.4.1 Similarity and response type

Figure 8-13 shows scatterplots for mean similarity (Sim(all)) and the three response type variables ($C_{(response)}$, $Nc_{(response)}$, $O_{(response)}$). Figure 8-14 shows the same data in the form of a bar chart arranged by the median similarity of the stimuli. Visual inspection of the scatterplots shows non-linear monotonic relationships between similarity and $C_{(response)}$ and $Nc_{(response)}$, but not with $O_{(response)}$. Thus, a Spearman's rank-order correlation was run to assess the relationship between mean rated similarity and the proportion of each response type. All tests are two-tailed.

- There was a statistically significant negative relationship between mean rated similarity and $C_{(response)}$, $r_s(28) = -.558$, $p = .001$
- There was a statistically significant positive relationship between mean rated similarity and $Nc_{(response)}$, $r_s(28) = .807$, $p < .001$
- There was no statistically significant correlation for mean rated similarity and $O_{(response)}$, $r_s(28) = -.236$, $p = .209$

These findings show that as similarity increases, designers are increasingly likely to produce non-combinational responses and increasingly unlikely to produce combinational responses. The finding that the proportion of 'other' codes is not statistically significantly correlated with similarity negates any issues with the trial difficulty confound associated with the proportion of 'O' codes (Section 8.1.3.2).

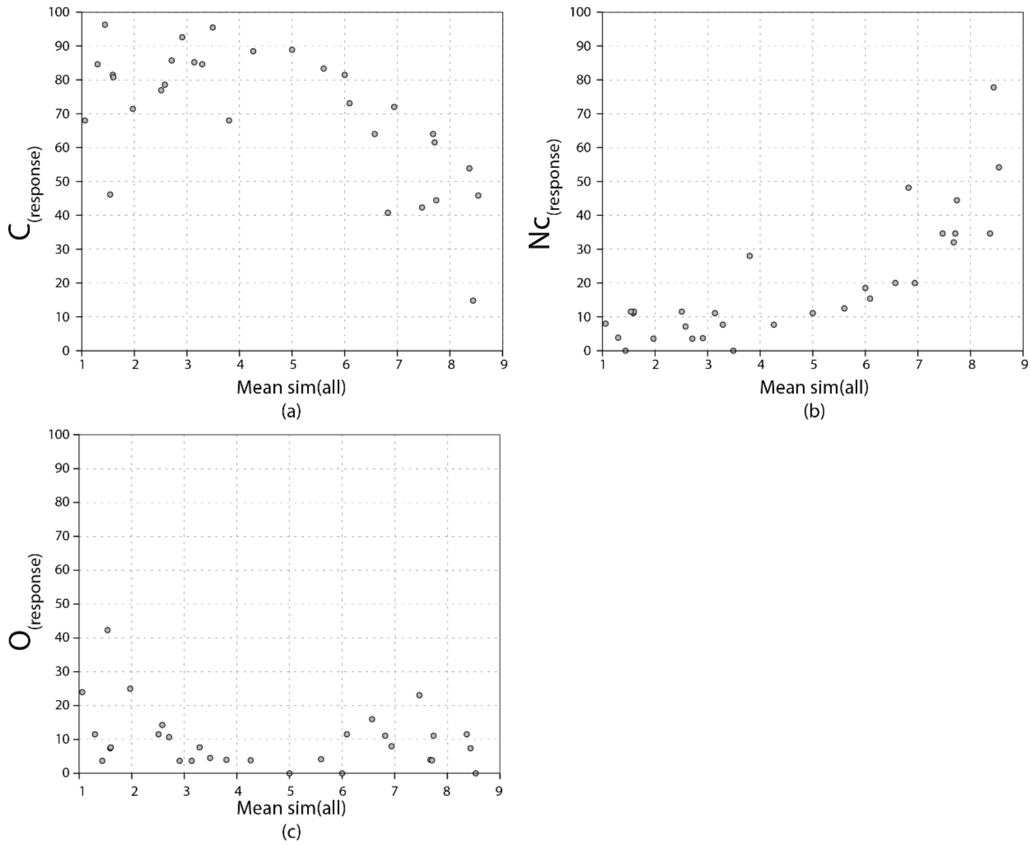


Figure 8-13 – Scatterplots showing mean similarity plotted against the proportion of response type for (a) $C_{(response)}$, (b) $Nc_{(response)}$ and (c) $O_{(response)}$ for each of the 30 stimuli.

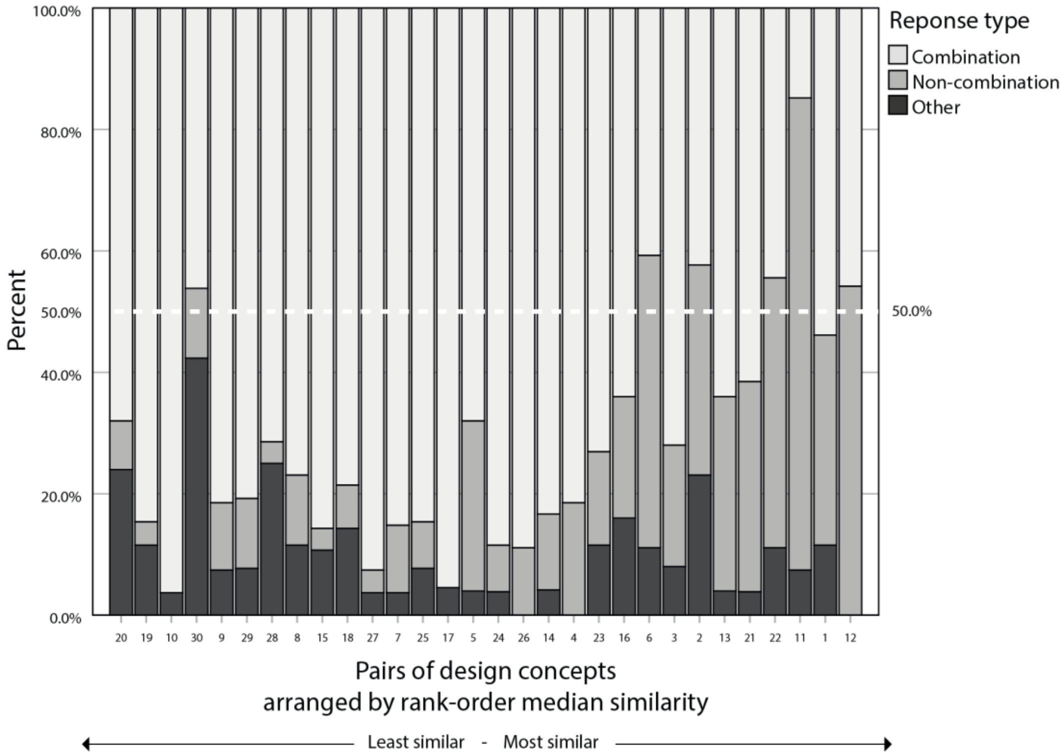


Figure 8-14 – Stacked bar chart showing $C_{(response)}$, $Nc_{(response)}$ and $O_{(response)}$ for each of the 30 stimuli, ordered by median similarity.

8.1.4.2 Similarity and combination type

Figure 8-15 shows scatterplots for mean similarity (Sim(all)) and the proportion of ideation type codes ($F_{(combo)}$, $R_{(combo)}$, $A_{(combo)}$). Figure 8-16 shows the same data in a stacked bar chart arranged by the median similarity of the stimuli. In this case, Nc responses were excluded, since they are not theoretically relevant to the proposed Dual-Process model of design concept combination.

Visual inspection of the scatter plots shows an approximately linear, negative relationship with $R_{(combo)}$ and no linear relationship with $A_{(combo)}$. The relationship between similarity and $F_{(combo)}$ is not clearly linear, but the data predominantly occupy the upper left triangle of the plot. Spearman's rank-order correlations were run to assess the relationship between mean rated similarity and the proportion of each combination type. All tests were two tailed.

- There was a statistically significant, moderate positive correlation between mean rated similarity and the proportion of featural combinations $F_{(combo)}$, $r_s(28) = .441$, $p = .015$
- There was a statistically significant, strong negative correlation between mean rated similarity and the proportion of relational combinations $R_{(combo)}$, $r_s(28) = -.592$, $p < .001$.
- There was no statistically significant correlation between mean rated similarity and the proportion of ambiguous combinations $A_{(combo)}$, $r_s(28) = .029$, $p = .878$.

These findings suggest that when designers combine pairs of design concepts, as similarity increases the proportion of featural combinations increases and the proportion of relational combinations decreases. The proportion of ambiguous combinations appears to be independent of base concept similarity. However, the statistical analyses should be interpreted with caution given the use of proportional data, the lack of clear linear relationships and the Likert-type scale data. The data are investigated at a more granular level in the next section.

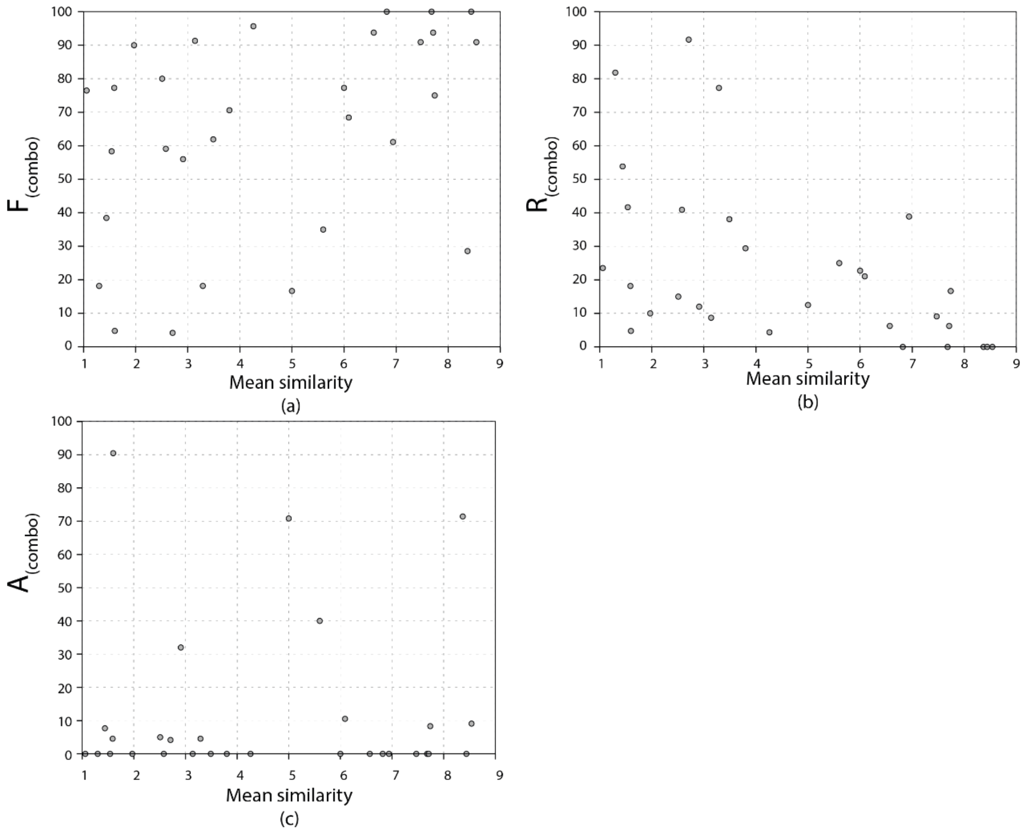


Figure 8-15 - Scatterplots showing mean similarity plotted against the proportion of combination type for (a) $F_{(combo)}$, (b) $R_{(combo)}$ and (c) $A_{(combo)}$ for each of the 30 stimuli.

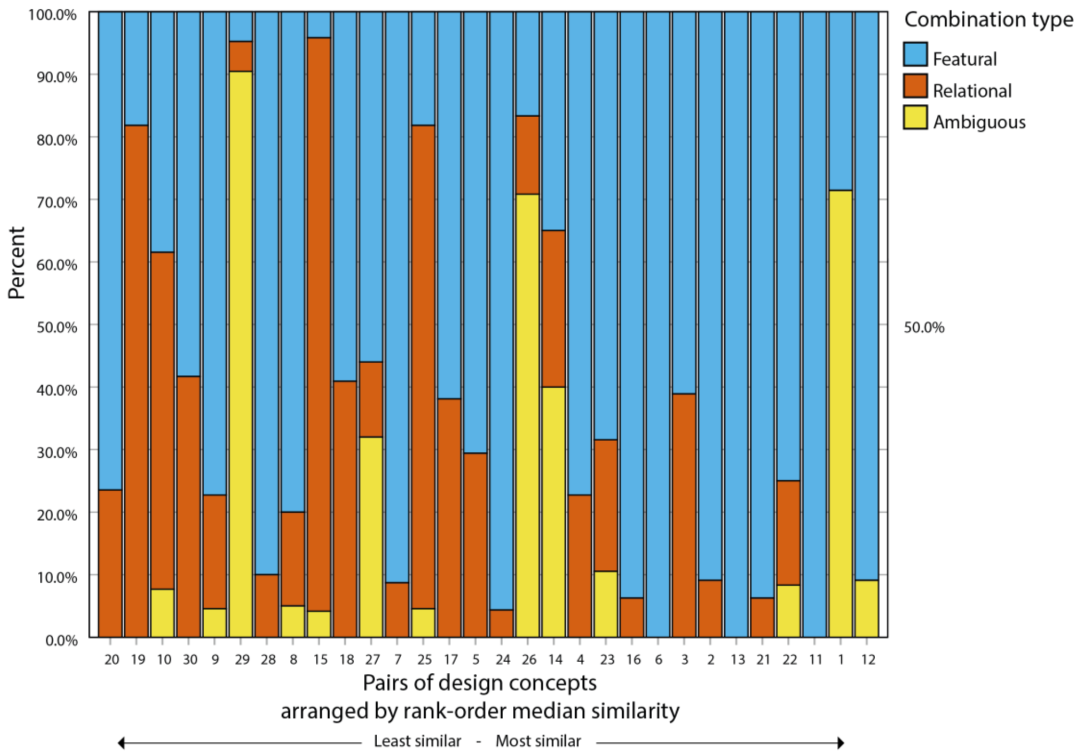


Figure 8-16 - Stacked bar chart showing the proportion of combination types for each of the 30 stimuli, ordered by median similarity.

When examined through the lens of the proposed Dual-Process model, the data in Figure 8-16 reveals three additional characteristics of the relationship between similarity and combination that may have meaningful theoretical implications.

Characteristic 1: The proportion of relational combinations appears to vary independently of similarity. Although there is a statistically significant, negative correlation between the similarity of a pair of design concepts and the proportion of relational combinations, stimuli of neighbouring rank-order similarity can have contrasting proportions of relational combinations. Examples are shown in Table 8-5 and Figure 8-17 and Figure 8-18. Each example lists two pairs of base concepts with the same value of median similarity alongside the proportion of ideation-type responses for that pair. Within each example, one pair has a relatively high proportion of Featural combinations and a low proportion of Relational combinations and vice versa for the other example. This suggests that the production of relational combinations is caused by something other than the similarity of the pairs of base concepts.

Characteristic 2: All responses have at least one Featural combination, but not all responses have a Relational combination. Pairs 12, 1, 11, 13 and 6 (Figure 8-16) received no Relational combinations.

Characteristic 3: Low proportion of relational combinations at high similarity. A second observation is that there are no relational combinations for the 3 most similar base concepts (Sc_12, 1 and 11).

Table 8-5 – Data for examples of pairs of design concepts with proximate similarity ratings and contrasting proportions of featural and relational combinations.

				Proportion of ideation type (%)			
Example	**Pair #**	**Sim (mean)**	**Sim (Mdn)**	**F**	**R**	**A**	**Nc**
1 (Same DT)	Sc_19 (DT06)	1.30	1	17.39	78.26	0.00	4.35
	Sc_20 (DT06)	1.06	1	68.42	21.05	0.00	10.53
2 (Different DT)	Sc_7 (DT03)	3.14	3	80.77	7.69	0.00	11.54
	Sc_25 (DT07)	3.29	3	16.67	70.83	4.17	8.33

DT = design task, Pair # denotes the base concepts and corresponds to Table A4-6, Mdn = median, Prop_x = the proportion of the specified variable to all four variables. Grey boxes highlight the contrasting proportions within each example.

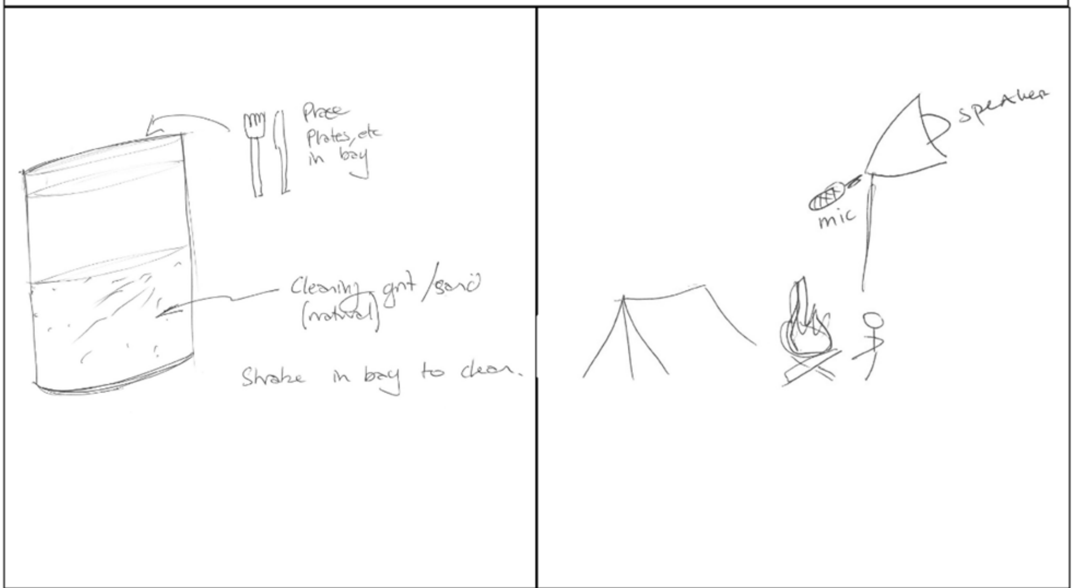
Pair 19, DT06
Median similarity = 1, high proportion of Relational combinations (78.26%)

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

Tent elevated off ground - less damage to plants on ground	Vacuum bag suction bin bag Put your rubbish in this bag, create a vaccuum, easy to carry rubbish around with you - light and small

Pair 20, DT06
Median similarity = 1, high proportion of Featural combinations (68.42%)

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



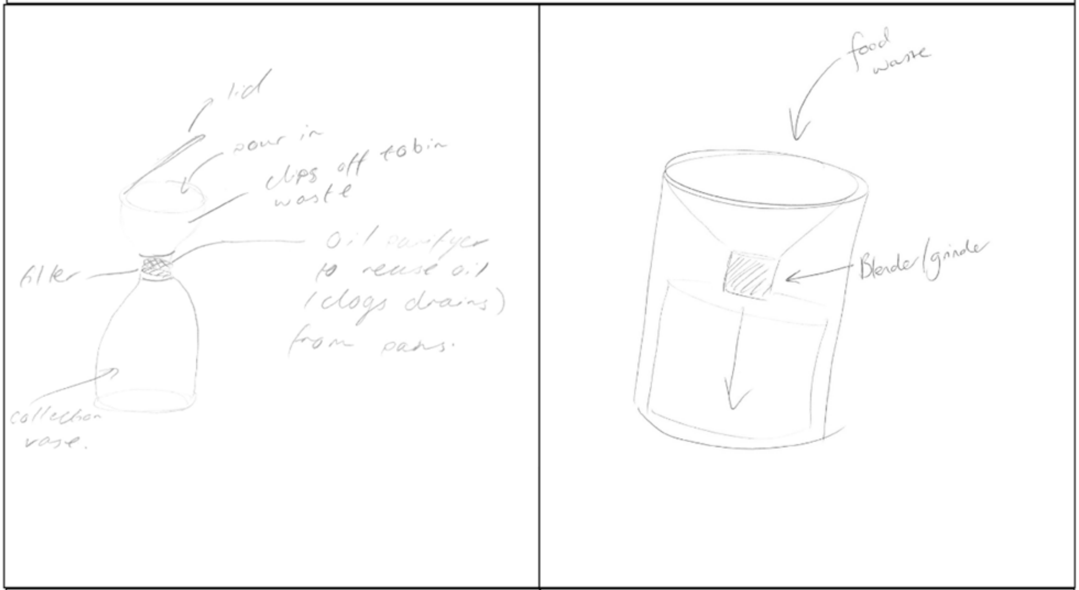
Shake to clean - no soaps!	the mic records all the noise you and your mates make @ campsite and then the speakers produce the mirrored wavelength to make it nice & quiet for the wildlife around

Figure 8-17 – Stimuli 19 and 20, examples of pairs of design concepts from the same design task with proximate similarity ratings and contrasting proportions of featural and relational combinations.

Pair 7, DT03

Median similarity = 3, high proportion of Featural combinations (80.77%)

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

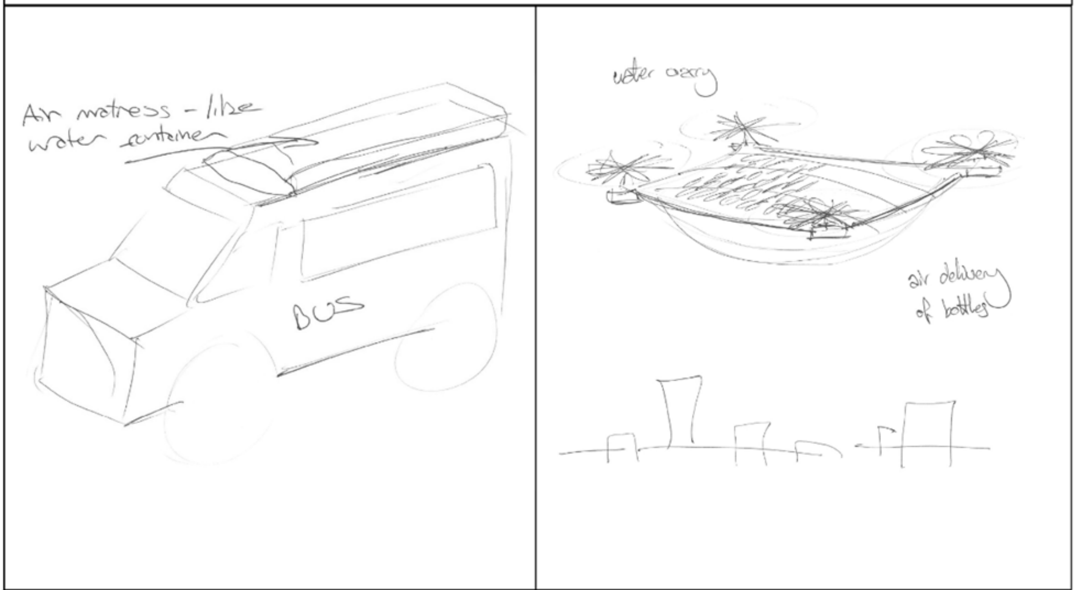


Oil is expensive and clogs drains. A filter means it could be reused	With the fod being grinded and blended it takes up less space and it is ready for recycling. People always like to have a gadget in their house too... even if its usefulness is questionable

Pair 25, DT07

Median similarity = 3, high proportion of Relational combinations (70.83%)

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Bus-mounted water transport	Heavy lift drones for mass delivery

Figure 8-18 - Stimuli 7 and 25, examples of pairs of design concepts from different design tasks with proximate similarity ratings and contrasting proportions of featural and relational combinations.

8.1.4.3 Combination and difficulty

Analyses were conducted to determine whether it was equally difficult to produce each type of combination.

Difficulty and response type. A Kruskal-Wallis H test was conducted to determine if there were differences in the mean ranks of the difficulty ratings between the three response types: combination, non-combination and other (Table 8-6, Figure 8-19). The distributions of difficulty ratings were not similar (Figure 8-19b) and so the Kruskal-Wallis H test is used to compare the mean ranks of the difficulty ratings. The difficulty ratings were statistically significantly different between the three response types, χ^2 (2) = 39.731, $p < .001$. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. There were statistically significant differences in difficulty ratings between O and C (p < .001), and O and Nc (p < .001), but not between C and Nc (p = 1.000). Thus, the trials in which participants produce 'other' responses are rated as being more difficult than those in which they successfully ideate. The production of combinational and non-combinational design concepts did not differ significantly in terms of difficulty.

Table 8-6 – Median and count values for rated difficulty across three design tasks.

	Median	Mean rank	n
(C) Combination	2	373.08	548
(Nc) Non-combination	2.5	372.61	152
(O) Other	4	537.79	78
Total			778

Note: Mdn = Median value, n = number of responses, DT = Design task.

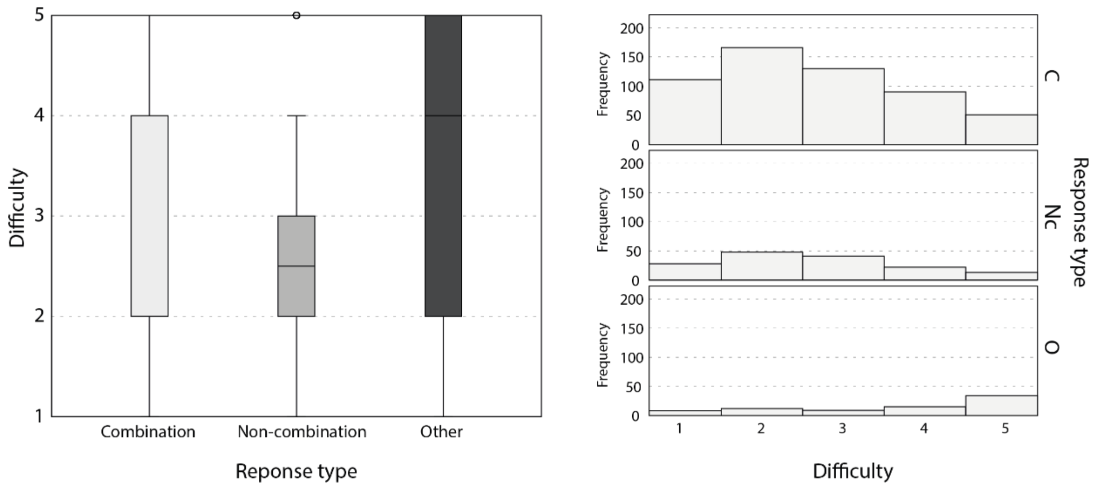


Figure 8-19 – Boxplots showing the difficulty ratings for each response type.

Difficulty and combination type for each trial. A Kruskal-Wallis H test was conducted to determine if there were differences in difficulty ratings between the four ideation types: featural (F), relational (R), ambiguous (A) and non-combinational (Nc) (Table 8-7, Figure 8-20). The shapes of the distributions of difficulty ratings were similar across the three response types (Figure 8-20b) and so the test was used to compare median values. The difficulty ratings were statistically significantly different between the four response types, χ^2 (2) = 17.954, $p < .001$. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. There were statistically significant differences between the difficulty ratings for Ambiguous combinations (*Mdn* = 2) and the three other ideation types: F (*Mdn* = 3) (p = .001), R (*Mdn* =3) (p = .001) and Nc (*Mdn* = 2.5) (p = .000), but not between any other pairs. Thus, when considering only successful instances of ideation, the trials in which participants produce ambiguous responses are rated as being less difficult than those in which they produce any other type of design concept. The production of F, R and Nc responses did not differ significantly in terms of difficulty.

Table 8-7 – Median and count values for rated difficulty across three design tasks.

	Median	Mean rank	n
(F) Featural	3	360.50	329
(R) Relational	3	372.56	147
(A) Ambiguous	2	260.08	72
(Nc) Non-combination	2.5	350.35	152
Total			700

Note: Mdn = Median value, n = number of responses, DT = Design task.

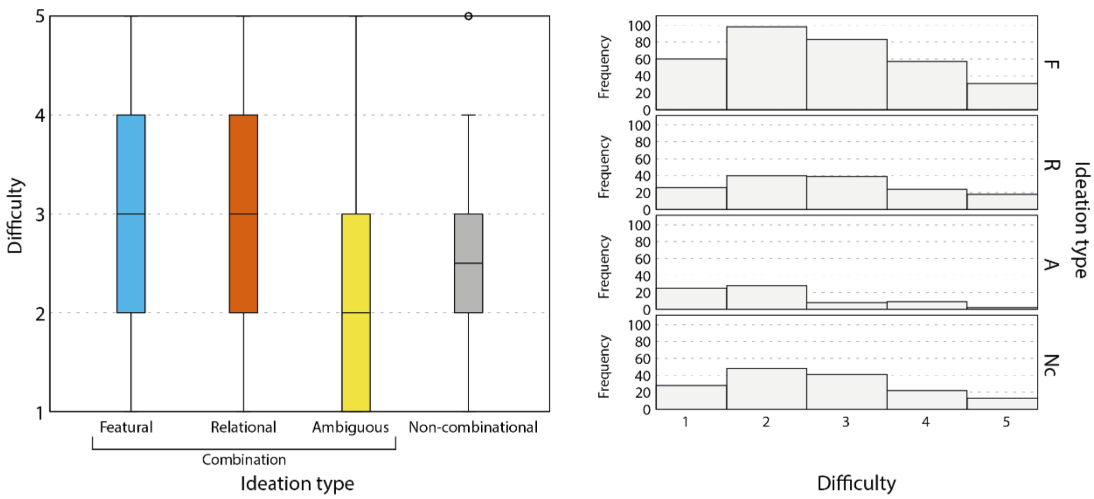


Figure 8-20 – Boxplots showing the difficulty ratings for each ideation type.

Summary. To provide additional insights into the difficulty of the combination trials, the differences in difficulty rating were investigated for the response types and ideation types.

- The tasks in which participants failed to ideate were rated as being statistically significantly more difficult than the others, but there were no significant differences in the difficulty ratings for C or Nc. This is consistent with the earlier finding that there was a positive correlation between trial difficulty and the number of O codes as a proportion of all responses (Section 8.1.3.2).
- When considering only successful ideation attempts (F, R, A, Nc), the production of ambiguous combinations (A) was statistically significantly less difficult than the production of featural combinations (F), relational combinations (R) or non-combinational responses (Nc).

8.1.5 Discussion of Combo-P3

Four research questions were tested to determine what kind of combinations designers produce when combining design concepts (RQ-C1), how prevalent each kind is (RQ-C2) and what the relationship is between these outputs and the similarity of the base concepts (RQ-C3) as well as the difficulty of combining them (RQ-C4).

8.1.5.1 The outputs of design concept combination

When tasked with combining pairs of design concepts, the designers predominantly combined their responses as expected (70.5% of all attempted trials). An unexpected finding was that they also produced design concepts that contain no evidence of combination. 19.6% of all attempted trials were coded as non-combinational responses. When designers *did* combine the base concepts, they produced Featural, Relational and Ambiguous combinations. This answers RQ-C1.

As shown in Table 8-3, featural combinations were the predominant form of combination, occurring in 60.3% of all instances of combination (332/551). Relational combinations were next at 26.7% (147/551) followed by ambiguous combinations at 13.1% (72/551). With reference to Figure 8-16, these ambiguous were predominantly associated with five base concepts, but at least one occurred for 13 pairs of base concepts. This answers RQ-C2, showing that featural combinations are the most prevalent, followed by relational and then ambiguous combinations.

RQ-C3 probed the relationship between similarity and combination type. Correlational analyses showed that similarity was positively correlated with the number of featural responses and negatively correlated with the number of relational combinations. There was no statistically significant correlation between similarity and the proportion of ambiguous combinations. However, the results of the correlational analyses should be interpreted with caution. As the correlations are conducted on proportional data, the positive and negative correlations are essentially two measures of the same change in proportion along the similarity scale. Moreover, visual inspection of the data shows that the relationship between similarity and the proportion of combination type is not fully captured by the aggregate correlations. It was found that (i) pairs of design concepts with near-identical similarity ratings can receive contrasting proportions of featural and relational combination, (ii) all pairs received at least one featural combination, but not all pairs received a relational combination, and (iii) there appears to be a cut-off of relational combinations at the highest levels of base-concept similarity. These three characteristics of the data were motivated by the expectations of the model proposed at the outset of the research, and so while they may more accurately characterise the relationship between similarity and combination type, they may also be influenced by expectation bias.

8.1.5.2 Planned hypothesis tests

The two combination model hypotheses (H-C1 and H-C2), proposed in Section 4.2.2.2, were not tested. Both hypotheses rely on the ability to distinguish between feature mapping and hybrid combinations. As noted in Section 8.1.5.2, it was not possible to distinguish between these two kinds of combinations, and so the planned hypothesis tests were not conducted.

The first hypothesis (H – C1) stated that relatively more similar pairs of design concepts should be associated with a greater proportion of hybrid combinations. This is an explicit, quantifiable prediction that stems from the first claim about similarity and combination type described This hypothesis was made redundant by the lack of a distinct 'hybrid' combination type.

H-C1 was the reason for testing the Structural Alignment model in similarity judgements in the first place. Initially, it was expected that there was a greater likelihood of finding support for Structural Alignment in similarity judgements than in combination, and that said support would act as a pseudo positive control to help make inferences about design concept combination. This too was rendered redundant by the lack of hybrid combinations.

The second hypothesis (H – C2) stated that there should be a correspondence between alignable differences and feature-mapping combinations, but not alignable differences and relational combinations. Specifically, the planned analysis for this hypothesis involved determining which *single* feature had been mapped and transferred from one design concept to another to facilitate the featural combination and to determine whether that single feature was alignable or nonalignable in the base concepts. Since it was not possible to isolate single features that had been transferred in design concept combination, this analysis could not be conducted.

8.2 Chapter summary

This chapter has presented the analysis and results of Combo-P3 to satisfy Obj. 4 – to *propose and test a cognitive model of design concept similarity judgements.* As outlined in the research plan (Figure 4-5), this was to be achieved by addressing three research questions and testing two hypotheses. The answers to the research questions are listed in Table 8-8. As discussed in Section 8.1.5.2, the planned hypotheses tests could not be conducted as they were rendered invalid by the results of the analyses conducted to address the research questions. The implications of these results for the proposed model of design concept combination are discussed in detailed in Chapter 9.

Table 8-8 - Summary of research questions, hypotheses and associated results for similarity experiments.

Ref.	Research question	Results
Combination model evaluation		
Research questions, computational level		
RQ – C1	What types of combinations do designers produce?	Response types – Designers responded to the combination task by combining (C), non-combinational ideation (Nc) or producing ‘other’ responses (O). Combination types - Designers produced featural (F), relational (R), ambiguous (A) and ‘other’ (O) combinations.
RQ – C2	What is the prevalence of each combination type?	Of the response types, designers most frequently combine the base concepts (70.5%), but also create non-combinational responses (19.6%) or fail to ideate (9.9%). Of the three types of combination, designers most frequently produce featural combinations (60.3%), followed by relational combinations (26.7%) and ambiguous combinations (13.1%).
RQ – C3	What is the relationship between concept pair similarity and the type of combined concept?	As similarity increases, the proportion of Nc responses increases and the proportion of C responses decreases. The tendency to produce ‘other’ combinations (i.e., fail to ideate) is not significantly associated with similarity. As similarity increases: the proportion of F combinations increases and the proportion of R combinations decreases. The proportion of A combinations is not significantly associated with similarity. Visual inspection of the data shows that (i) the proportion of R combinations can vary substantially for pairs of stimuli of proximate similarity, (ii) all pairs have an F combination but not an R combination, and (iii) very high similarity concepts have few or no R combinations.
RQ – C4	What is the relationship between concept pair combination difficulty and the type of combined concept?	The relationship between similarity and combination type is not confounded with trial difficulty. Difficulty is positively correlated with the proportion of $O_{(response)}$ combinations but with no other outcome measures. When considering response types, the trials in which participants produce ‘other’ responses are rated as being more difficult than those in which they successfully ideate. The production of C and Nc design concepts did not differ significantly in terms of difficulty. When considering only successful ideation attempts, the production of A combinations was statistically significantly less difficult than the production of F, R or Nc design concepts.

9 DISCUSSIONS

This chapter presents three discussions. The results of the experiments conducted to test the similarity and combination models are discussed in Sections 9.1 and 9.2 respectively. A general discussion of the thesis as a whole is presented in Section 9.3.

Whereas previous discussions have addressed the results of individual experiments (Sections 7.1.2, 7.2.5, and 8.1.5), the discussions in Sections 9.1 and 9.2 provide an assessment of the evidence for or against each model based on the experimental results. Inferences drawn about designer cognitive processing are discussed in terms of their consistency with prior research, their generalisability, and implications for design research and practice. Recommendations are provided for conducting further research into similarity judgements and combination cognitive processing in design.

The general discussion in Section 9.3 addresses objective 5 by critiquing the research in the thesis as a whole to identify strengths, limitations and areas for future work.

9.1 The cognitive processes of design concept similarity judgments

The second objective for the research in this thesis (Obj. 2) was to propose and test a model of design concept similarity judgements. The purpose of this was to gain knowledge about how designers make similarity judgements of early-stage design concepts. This was done, in part, to aid in generating inferences about design concept combination and thus support the parallel stream of research into combination cognitive processes. It was proposed that designer similarity judgements of design concepts occurred via two processes of comparison and scenario creation (Section 4.1). It was further proposed that the comparison process operated via a structural alignment algorithm as described by the Structural Alignment model of similarity judgements (Gentner and Markman, 1994, 1997; Markman and Gentner, 1993b, 1993a).

9.1.1 Evaluation of the proposed model

The results of Sim-P2 (Section 7.1) Sim-P3 (Section 7.2) and Combo-P3 (Section 0) are consistent with a single-process, comparison-based model of similarity judgements. There is no evidence in support of a scenario creation process that operates on thematic relations between design concepts. Although it was proposed that a comparison-based similarity judgement process would operate via a process of Structural Alignment, the proposed model was falsified and no claims can be made about algorithmic-level processing.

9.1.1.1 Evidence for a single-process of comparison

Evidence in support of a comparison-based similarity process comes from (i) the predominance of featural combinations found in the similarity explanation task, (ii) the finding that similarity can be predicted as a function of commonalities and differences, and (iii) the ruling out of a stimulus compatibility confound.

Initial evidence for a comparison-only model of similarity judgements came from the predominance of feature-based explanations for similarity judgements (Sim-P2). The participants' responses were overwhelmingly featural and no responses contained clear evidence of thematic processing. The lack of thematic explanations in Sim-P2 is in stark contrast to the prevalence of thematic processing identified in linguistic conceptual combination research. Wisniewski and Bassok (1999) found that 89% of participants (57/64) produced at least one thematic explanation for a similarity judgement, with a median of 4 per person, and 38 of the 47 object pairs used as stimuli received thematic explanations. In Sim-P2, none of the 11 participants produced a clearly thematic explanation for similarity, and none of the pairs of design concepts received any clearly thematic explanations. The data thus provide support for a single, feature-based process of similarity judgements. However, this finding does not disprove the existence of thematic responses in design concepts, as thematic responses may arise with larger samples or a different experimental measure (see a further discussion in Section 9.1.2)).

The test of the Structural Alignment model provides further support for a comparison process. In any comparison-based model of similarity judgements, similarity should be a function of the commonalities and differences listed for a pair of concepts, and as similarity increases so too should the number of commonalities. H1a and H2 set out to test this basic prediction. When designers make similarity judgements of design concepts, the relative similarity of a pair of design concepts can be predicted by the number of commonalities and differences listed for the pair (explaining 73.9% of the variance) with commonalities contributing more to similarity than differences detract from it (H1a). Further, as the relative similarity of a pair of design concepts increases, so too does the number of commonalities that those concepts have (H2), as shown by correlational analyses. These provide evidence that similarity judgements are carried out by a comparison process of some kind, where similarity is a function of commonalities and differences, but they do not provide specific support for the Structural Alignment *per se.*

The findings from Combo-P3 can further support the claim for a comparison process by ruling out a potential stimulus compatibility confound. One possible criticism of the similarity explanation task (Sim-P2) is that the predominance of feature-based explanations could be the result of a stimulus compatibility confound. Introduced in Section 3.2.3, stimulus compatibility is the phenomenon identified in conceptual combination wherein pairs of concepts can be more or less compatible with featural or thematic processing depending on the extent to which they have shared features and whether a plausible external relation exists between them (Wisniewski and Bassok, 1999). It could be argued that the pairs of design concepts used in Sim-P2 were created in such a way that made them predominantly compatible with featural, but not thematic processing. In turn, this might imply that thematic processing could arise in similarity judgements of different stimuli.

A stimulus compatibility confound can be ruled out because the same design concepts that received exclusively featural explanations in the explanation task (Sim-P2) were frequently processed via scenario creation to produce relational combinations in Combo-P3. As shown in Table A4-5 the stimuli set used in the explanation task (Stimuli Set B) shared 18/30 stimuli with Stimuli Set C used in the design concept combination task [17]. The top row of Table 9-1 lists the 18 pairs of design concepts used in both studies, all of which received exclusively featural or 'other' explanations. The row below (R) lists the number of relational combinations that were produced for the same pairs in Combo-P3. 12/18 elicited at least one relational combination with one pair of concepts eliciting 78.6% relational responses. This shows that it is possible to place the base concepts used in the similarity explanation task into scenarios, but that designers do not do so during similarity judgements. Thus, the predominance of featural explanations found in the similarity explanation task was likely not a consequence of a stimulus compatibility confound.

An illustrative example is provided in the case of stimulus pair Sc_C25 (Figure 8-1). All 9 explanations provided for this pair were based on featural justifications such as: "*Both Cary water, both are transport, but do so by different means*", and "*Both are sheets for carrying water*" (Table 9-2, below). In the combination task, however, 17 of the 26 combined concepts created from this pair were made by placing the two design concepts into scenarios like the example of a relational combination in Figure 8-3.

[17] It would have been better to compare two studies from the same phase, but since the explanation task was only used in phase 2 (Sim-P2) and no combination experiment was conducted in this phase, this is not possible. Nonetheless, these two studies share enough common stimuli to facilitate some comparison because stimuli set 3 was made by modifying parts of stimuli set 2.

Table 9-1 - Relational combinations in stimuli used in Comb-P3 that were also used in Sim-P2

Pair	**1**	**2**	**6**	**7**	**10**	**11**	**12**	**13**	**14**	**15**	**18**	**21**	**22**	**23**	**25**	**28**	**29**	**30**
R	0	1	0	2	14	0	0	0	5	22	9	1	2	4	17	2	1	5
%	0	4	0	7	52	0	0	0	21	79	32	4	7	15	65	7	4	19
n	26	26	27	27	27	27	24	25	24	28	28	26	27	26	26	28	26	26

Pair = stimulus pair number, R = number of relational combinations, % = relational combinations as a percentage of n to zero decimal places, n = number of attempted trials for a given pair in Combo-P3

Table 9-2 – Similarity explanations for base concepts Sc_25

#	**Similarity explanation**
1	Water transport vehicle only real similarity.
2	both high tech, one flys other drives
3	Both Cary water, both are transport, but do so by different means.
4	water transportation
5	Only one similarity in that both allow mass water transportation using a vehicle. But major differences - air transportation vs land transportation, one's water transportation feature is a secondary quality of the products primary purpose (to transport people).
6	Both are sheets for carrying water
7	Completely different means of carrying water. One is a large water container attached to existing methods of transport. The other is specially designed infrastructure to carry pre-packaged water bottles.
8	different transport methods and cargo (Water and Bottles)
9	Concepts vary in the technology used, the method (land/air) and the manner of storage.
10	It's not completely clear what the water is placed in in the second design i.e. what are the water containers? Therefore hard to tell if that part of the design is similar. Similarity is that vehicles are used to transport water, but vehicles are different.
11	-No response-

9.1.1.2 The role of thematic processing in comparison and similarity

Ruling out a stimulus compatibility confound not only strengthens the evidence *for* a comparison process but also provides evidence *against* a thematic similarity process. If designers can place design concepts into scenarios during combination, but they don't do it during similarity judgements, then this suggests that designers do not make similarity judgements based on thematic (external and complementary) relations between the design concepts.

Although there was no evidence for a thematic similarity judgement process, there were unexpected instances of thematic processing in the commonality and difference listing task (Sim-P3), raising questions about the role of thematic processing in comparison processes more generally. In the process of coding differences as ADs or NDs, two of the 1618 differences listed were identified as being thematic by the researcher despite not being part of the coding scheme. When asked to list differences for Sc_20 (Figure 9-1), one participant provided two *negations* of thematic relations: *"Mic/speaker would not enable the cleaning of cutlery"* and "*Bag would not cause you and your mates to be quiet."* Both responses are examples of thematic relations since the participant has mentioned external and complementary relations between one design concept and the other. This indicates that some thematic processing occurred during a comparison task at a very low prevalence (0.12%). This was not expected since the task was assumed to elicit comparisons and not scenario creation processing. Revisiting the literature revealed that similar spontaneous listing of thematic relations in a commonality and difference listing task had been observed previously when listing commonalities (Markman and Gentner, 1993a, p.522), but the prevalence of these responses was not reported.

Overall, the lack of thematic explanations in the similarity-explanation task and the low prevalence of thematic processing on the commonality and difference listing task is consistent with the view that thematic processing may occasionally intrude on, comparisons (Gentner and Brem, 1999; Honke and Kurtz, 2019). The results of Sim-P2 indicate a predominantly featural model of similarity judgements, but the influence of thematic relations on similarity judgements cannot be ruled out completely. Future experiments would benefit from including measures to detect the potential influence of thematic processing in case it is more prominent with different study characteristics.

The contrast between the present research (finding no thematic explanations) and Wisniewski and Bassok (1999) may be attributed to two key differences between the studies: the design experience of the participants and the kinds of concepts being judged. Product design engineers may generally have a predisposition to evaluating the common and different features of design concepts owing to training in methods that prescribe composition or decomposition. All participants had completed at least 1.5 years of education in a PDE degree at university level, during which time they had training in functional decomposition and morphological composition methods. Design concepts may also not elicit thematic associations in the same way that category concepts do. When presented with category concepts like COW-MILK, the simultaneous activation of knowledge schema for cows and milk can lead to the recognition of the thematic relation that cows *produce* milk. In contrast, design concepts are inherently novel representations of objects that do not yet exist, meaning the person making the similarity judgement will not have built up knowledge of thematic associations between any two design concepts.

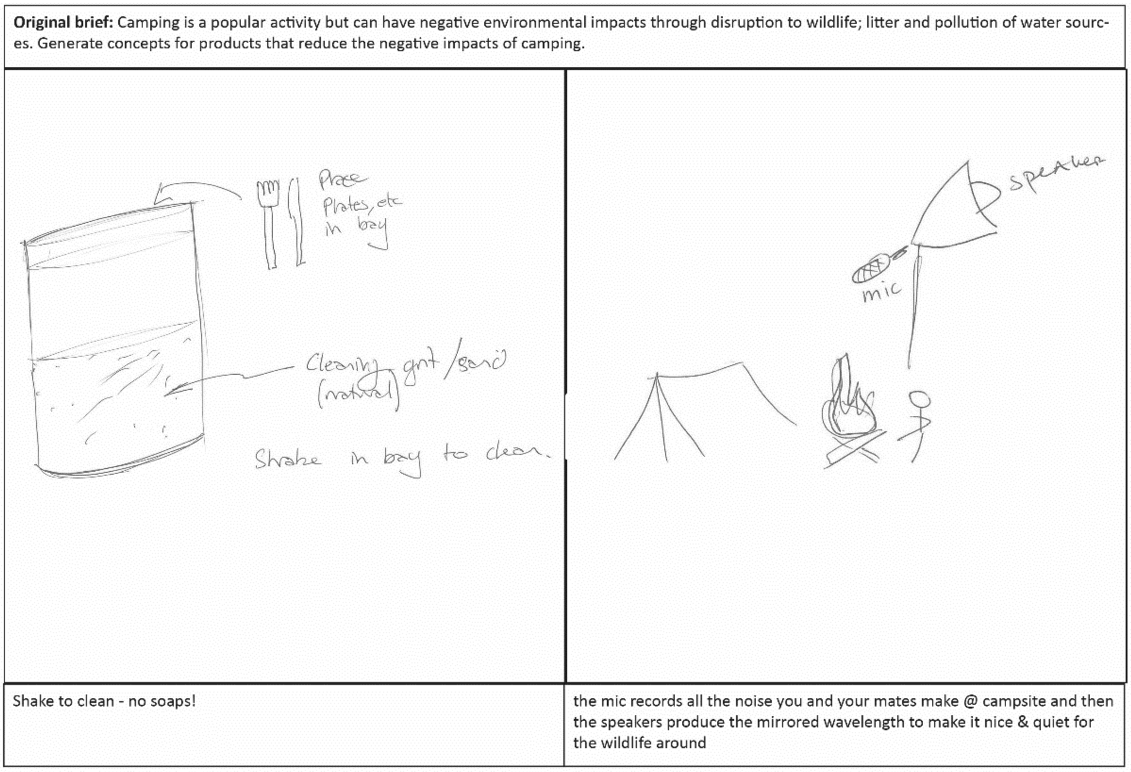


Figure 9-1 – Base concept pair Sc_20

9.1.1.3 Rejection of the Structural Alignment model

As discussed in Section 7.2.5, one of the hypotheses of the proposed Structural Alignment model was falsified. It was expected that there would be a positive correlation between the number of commonalities that participants list for a pair of design concepts and the number of alignable differences that they list. In a Structural Alignment view of similarity, a designer's perception of similarity would be a product of the mental alignment of the structured representations of the items being combined. The similarity process would seek to align the relational structures of the design concepts whilst attempting to maximise the correspondence between higher-order causal relations. As the number of commonalities increased, so too then should the number of alignable differences (the differences related to the common relational structure). Contrary to expectations, in Sim-P3 there was a non-significant, negative correlation between the number of listed commonalities and differences.

The falsification of H3 does not mean that the Structural Alignment model has been conclusively rejected. To reject the model outright would be an example of naïve falsificationism. However, as stated in Section 1.3.1, the research approach adopts the assumption of epistemological fallibilism, meaning that inferences can be demonstrated to be wrong. As discussed in Section 7.2.5, it may be possible to find additional support for the Structural Alignment model by finding alternative explanations for the falsification of H3. Nonetheless, at present, no claims can be made about how comparison-based similarity judgements of design concepts would operate.

9.1.2 Consistency with prior research in design

As stated in Section 3.1.3, no research has been conducted with the explicit aim of modelling the cognitive processes involved in designer similarity judgements generally, or of design concepts in particular. The closest relevant research involves the comparison of human similarity judgements with quantitative measures of product similarity (Nandy and Goucher-Lambert, 2022). In this study, human similarity judgements were taken from forced-choice triplet queries (asking whether A is more similar to B or C) and compared with six quantitative measures taken directly from product function models. Three of the quantitative measures were feature-matching measures (Simple Matching, Cosine and Jaccard coefficients), and the other three were network similarity measures. Notably, the Jaccard coefficient is a limited form of the Contrast model formula (Equation 1) for symmetrical similarity (i.e., when [a] and [b] = 1). Thus, this study provides a measure of association between a quantitative measure of similarity computed via the Contrast model and human similarity judgements. This is relevant because the Contrast model is, like the Structural Alignment model, a feature-based model that assumes similarity judgements occur via a process of comparison.

It might have been expected that if the participants made similarity judgements via a process of comparison, the Jaccard coefficient would more closely align to the human similarity judgements than the other measures. The results showed relatively homogeneous alignment between the quantitative measures and human similarity judgements. No single measure matched best with human ratings across both levels of abstraction, but some showed statistically significantly improved performance in specific circumstances, such as at individual levels of abstraction. The Jaccard Coefficient was not the best reflection of the human similarity judgements.

One reason for the relatively homogeneous performance may be that all six quantitative measures that were used were derived from the same functional flow representations. Products such as 'VHS player' or 'stapler' were represented in terms of functional flows, and those functional models were represented as binary matrices. Products either had a specific function (1) or did not have that function (0). On the other hand, the humans making the triplet judgements were presented with the same functional flow data in descriptive text form but also with titles and pictures of each product. Functional representations likely capture only a partial portion of a person's mental model of a product, ignoring other facets such as structure and behaviour that could influence the human similarity judgements. The quantitative measures may have been constrained in the extent to which they could fully align with human similarity judgements, leading to the relatively homogeneous ratings from the quantitative similarity measures.

9.1.3 Validity

Following the evaluation of the proposed similarity model, it may be claimed that design students make similarity judgements of early-stage, sketch-based design concepts via a process of comparison. Threats to statistical conclusion and internal validity have been addressed in the discussion of results (Sections 7.1.2 and 7.2.5) and the discussion of the model (Section 9.1.1). The generalisability of the claim is dependent on the strengths and weaknesses relating to construct, external and ecological validity.

Four of the study characteristics (participants, stimuli, IV manipulation and experimental setting) were shared by research in the same phase, and so are common to both the similarity and combination research in this thesis. At present, the claim that designers make similarity judgements of design concepts is limited to design students from two universities in Scotland. The knowledge claim may generalise to other product design engineering students in the United Kingdom, but it is not clear to what extent demographic differences, educational experience or individual differences in cognitive processing would influence similarity judgements of design concepts. Likewise, the use of multiple design tasks suggests that the findings may generalise to new design tasks. However, it should be noted that one design tasks was dropped between stimuli sets B and C (Section 6.4) because they failed the similarity manipulation check, suggesting that some design tasks may elicit different kinds of processing from designers.

Two issues of validity that are specific to the similarity research pertain to the measurement of similarity and the generalisability of findings to other kinds of design concepts. A unique strength of the similarity research is the use of two measures of similarity (similarity ratings and explanations), thereby avoiding a mono-operation bias. However, the research does suffer from a mono-method bias and could be further strengthened by using different tasks. A comparison-based model should be able to make predictions about task performance in other operationalisations of similarity, such as the speeded difference task (Gentner and Gunn, 2001; Gentner and Markman, 1994) or a forced-choice task (Markman and Gentner, 1996). A limitation of similarity judgements over other tasks is that while they are useful for measuring the extent to which two things are similar, they do not provide any direct insights into which specific features or dimensions the constituent concepts are comprised of and which of those features or dimensions contribute to a sense of similarity (Hebart *et al.*, 2020). With regards to the measurement of comparison, the measure of alignable and nonalignable differences used in this thesis is superior to the only other comparable measure in design (Nagai *et al.*, 2009). ADs and NDs were coded directly from a commonality and difference listing task and so the measure directly reflects the differences that designers perceived between the concepts. On the contrary Nagai *et al.* (2009) did not elicit a comparison process from their participants. To code commonalities and differences, they asked participants to list features of the design concepts produced through combination, and then coded those features. Thus, their ADs and NDs seem to pertain to the triad of the base concepts *and* the output. This is an entirely different measure than that used in the original conceptual combination literature and in this thesis.

A further potential issue is that models of similarity tend to only be compatible with specific kinds of stimuli (Goldstone and Son, 2012). The authors note that “researchers tend to consider, or in many cases create, entities to be compared that are compatible with their proposed model.” (p.22-23). They also note, citing Hodgetts *et al.* (2009), that “Cases in which a researcher proposing one model is able to accommodate the stimuli used by another modelling tradition are impressive but rare” (p.23). Other cognitive models may be more applicable for different kinds of design concepts. For example, Chaudhari *et al.* (2019) propose a knowledge modelling approach, where similarity is a function of the extent to which one can generalise from one artefact to another, accounting for design parameters, performance parameters and knowledge of the mapping between the two sets of parameters. Their method highlights the increasing importance of scientific knowledge used in later stages of the design process such as embodiment design. Future research could explore the generalisability of a comparison model of design concept similarity judgements across a range of different representations.

9.1.4 Implications

The main implication of the similarity research pertains to the proposed Dual-Process model of design concept combination. The results of the similarity experiments were important for the combination experiments in two ways. First, the test of the Structural Alignment model of similarity judgements provided a more robust test of structural alignment processing of design concepts. 'More robust' because the Structural Alignment model of similarity judgements makes more predictions and has received a greater deal of empirical support than the proposal that structural alignment is involved in conceptual combination. Secondly, the Structural Alignment model of similarity judgements was the foundation for Wisniewski's claims about Structural Alignment processing in conceptual combination. Thus, the falsification of the Structural Alignment model in similarity judgements implies that there may be something unique about cognitive processes by designers or with design concepts and that this might also mean that the claims about combination cognitive processing will not be borne out.

Aside from the implications for the combination studies, the findings have implications for more research into designer similarity judgements. The knowledge that designers make similarity judgements via comparison could open up new areas for cognitive process research of other evaluative or comparative processes. For example, similarity has been implicated in novelty ratings (Brown, 2016). If one could, e.g., control for the effect of similarity on expert novelty ratings then it may be possible to delineate other components of novelty judgements.

In practice, designers may benefit from an awareness of their biases in similarity judgements. For example, it was found that commonalities influence similarity more than differences. Thus, when making similarity judgements, or when carrying out processes that involve the same underlying processes as similarity judgements, designers may have a commonality bias that could be unknowingly influencing their thought processes.

Knowledge of how designers make similarity judgements may be used in retrieval-based design systems, such as analogical or case-based reasoning systems, or combinational tools that retrieve and recommend concepts for combination. Many retrieval systems use spatial 'distance' based measures of similarity, but it may be more appropriate to understand the exact features that drive similarity perceptions and to retrieve analogues or inspiration based on common and different features, rather than a less informative measure of distance.

9.1.5 Future work: towards an algorithmic level model of design concept similarity judgements

There are two broad directions for future research towards an algorithmic-level model of design concept similarity judgements. The first would be to probe the lack of support for H2 through additional analyses or experiments. Two hypotheses are provided that would facilitate further testing of a Structural Alignment model. They are based on the discussion in Section 7.2.5 in which it was proposed that the expected relationship between commonalities and alignable differences could be masked if participants were listing surplus ADs for low similarity pairs or a lack of ADs for high similarity concepts.

$H_{(N1)}$ – Numerous alignable differences exist, but only a limited number are being listed. Participants are not listing all the alignable differences that could conceivably exist between any two design concepts, especially for high similarity pairs. This hypothesis can be tested by analysing the *variety* of the differences listed by participants. Gentner and Markman (1994) found that more similar concepts had a greater variety of differences associated with them and less similar concepts had a lesser variety of responses. The increased variety of responses is to be expected in a structural alignment view of similarity since the increasing number of commonalities (overlap in relational structure) should lead to an increasing number and variety of alignable differences. If participants listed a greater variety of differences for similar design concepts than for dissimilar ones, this would suggest that an increased number of alignable differences exist between any two given concepts but are not being listed.

$H_{(N2)}$ There are surplus ADs for low similarity pairs. Participants may find it just as easy to list alignable differences for low commonality pairs as they do for high similarity pairs. For example, it may always be possible to identify alignable differences in e.g., the behaviour of a pair of artefacts along dimensions like weight, cost or simplicity. If this were the case, one would expect to see an increased number of these behaviour differences for pairs that are relatively dissimilar and have few commonalities. This could be tested by coding the types of alignable differences and looking for an association between the type of AD and the number of commonalities listed for a pair.

If tests of these hypotheses failed to find support for the Structural Alignment model, future research could then take one of two paths. It might be the case that designers do make similarity judgements of early stage design concepts via an alignment process, but that some aspects of the designer's experience or the design concepts mean that the alignment process manifests in unexpected ways. This would require a reconceptualization of Structural Alignment and the proposal and testing of new predictions. Alternatively, other models may provide more appropriate representations of design concept combination. The transformation model proposed by Chaudhari *et al.* (2019) or the variety of quantitative models tested by Nandy *et al.* (2020, 2022), Nandy and Goucher-Lambert (2021), may provide the basis for cognitive models of design cognition.

9.2 The cognitive processes of design concept combination

The aim of the research in this thesis was to model the cognitive process(es) of combination in conceptual product design engineering. The focus of the research was on *design concept* combination i.e., the combination of past solutions to a design problem to create new design concepts that address the same design problem. Towards this aim, Objective 3 was to propose and test a cognitive model of design concept combination. To achieve this objective, it was proposed that product design engineers combine design concepts via dual processes of (i) comparison and construction and (ii) scenario creation. It was further proposed that the comparison aspect of the comparison and construction process operates via the same structural alignment algorithm thought to be involved in design concept similarity judgements (Section 3.2). Research questions were addressed to gain an initial impression of whether design concept similarity judgements occurred via the proposed dual processes. If it appeared that they did, then hypotheses would be tested to determine whether the comparison and construction processes operated via a structural alignment algorithm.

9.2.1 Evaluation of the model

The results from Combo-P3 partially align with the expectations of the proposed model, but the model did not account for all of the findings. The proposed model can be evaluated in three parts: (i) the implications of the non-combinational responses, (ii) the extent to which the model successfully described the outputs of design concept combination, and (iii) the extent to which the model successfully described the relationship between similarity and combination type.

9.2.1.1 The implications of non-combinational responses

The finding that designers produce non-combinational responses when asked to combine pairs of design concepts was not accounted for by the proposed model. One explanation for these responses is that they resulted from distinct cognitive processes other than combination, meaning they can be effectively excluded from discussions of the 'genuine' combinational responses. A second explanation is that combination *did* occur, but that it was not detectable by the independent raters and so the analyses missed a large proportion of genuine combinational responses. Examination of the non-combinational responses and their relationship with similarity can reveal the circumstances in which they occur and support inferences about why they occur.

Although the non-combination responses were not coded into different types, it may be inferred that they are the result of other ideation processes such as *generation* and *transformation.* Two examples are shown of responses to the base concepts SC_13 (Figure 9-2).

In some instances, participants took features common to both concepts and used them as the basis of a new concept, or took features of a single concept and used them as the basis of a new concept. This can be construed as examples of *transformation* (Table 1-2) or what other authors have called modification (Daly *et al.*, 2012; Goel, 1995). An example is shown in Figure 9-3. It can be inferred that the concept is a transformation of the right-side base concept as none of the features of the left-side concept can be identified in it.

In other responses, participants produced design concepts that had no commonalities with either base concept other than perhaps having some relevance to the original brief (Figure 9-4. This can be construed as examples of *generation* (Table 1-2). In this example, the participant has created a new product that could be used as an accessory with either of the base concepts. This could be construed as a form of generative scenario creation, where a newly generated product forms a thematic relation with one of the base concepts. It does not fit the definition of a relational combination which requires that an external relation is placed between the two base concepts.

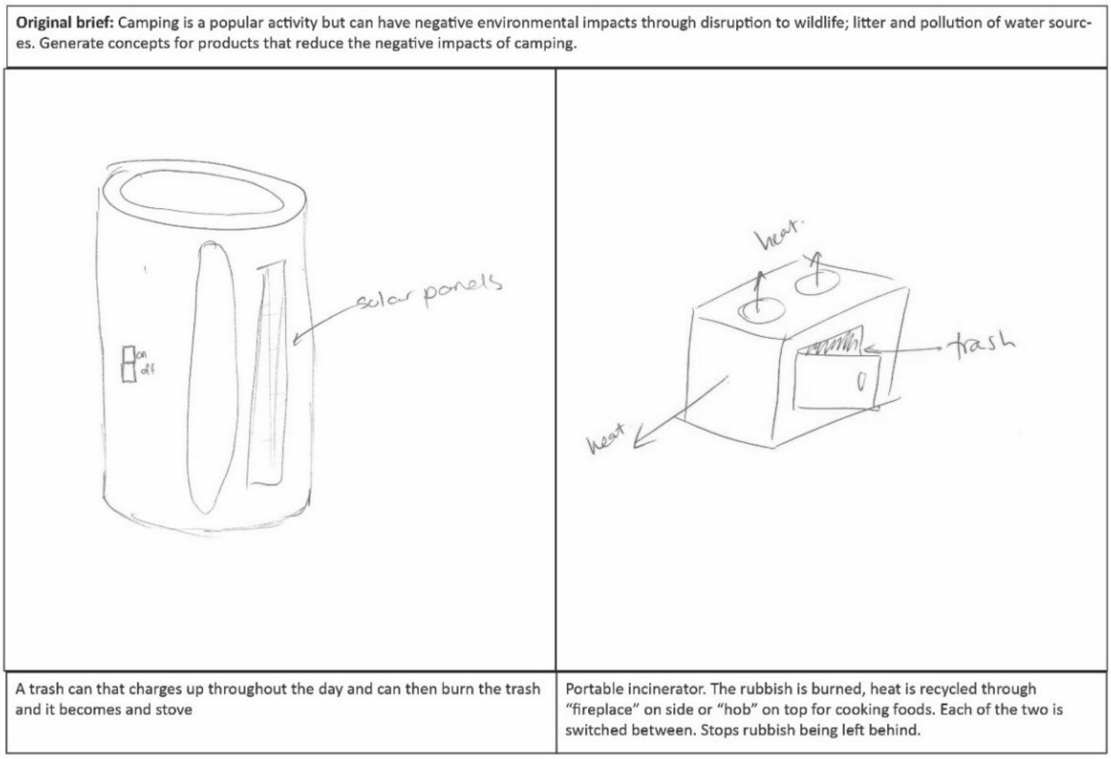


Figure 9-2 - Base concepts SC_13

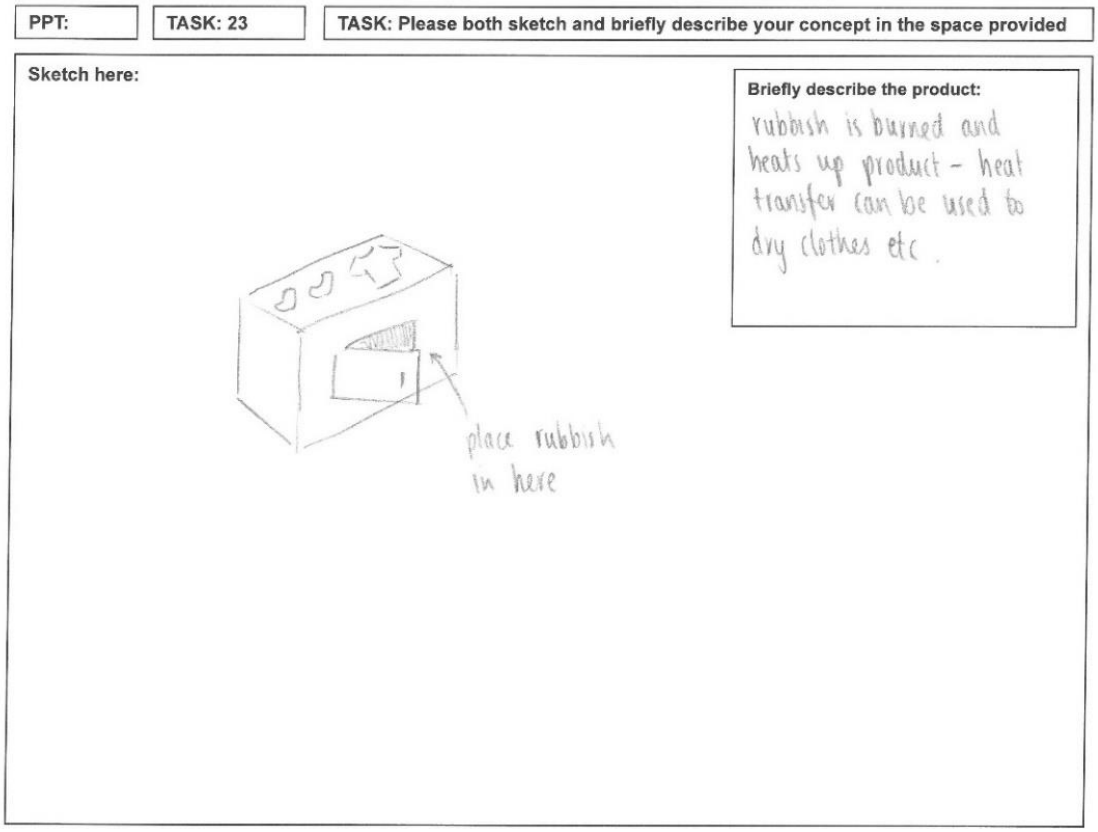


Figure 9-3 – An example of a non-combinational, 'transformation' response. The description reads: "Rubbish is burned and heats up product – heat transfer can be used to dry clothes etc.".

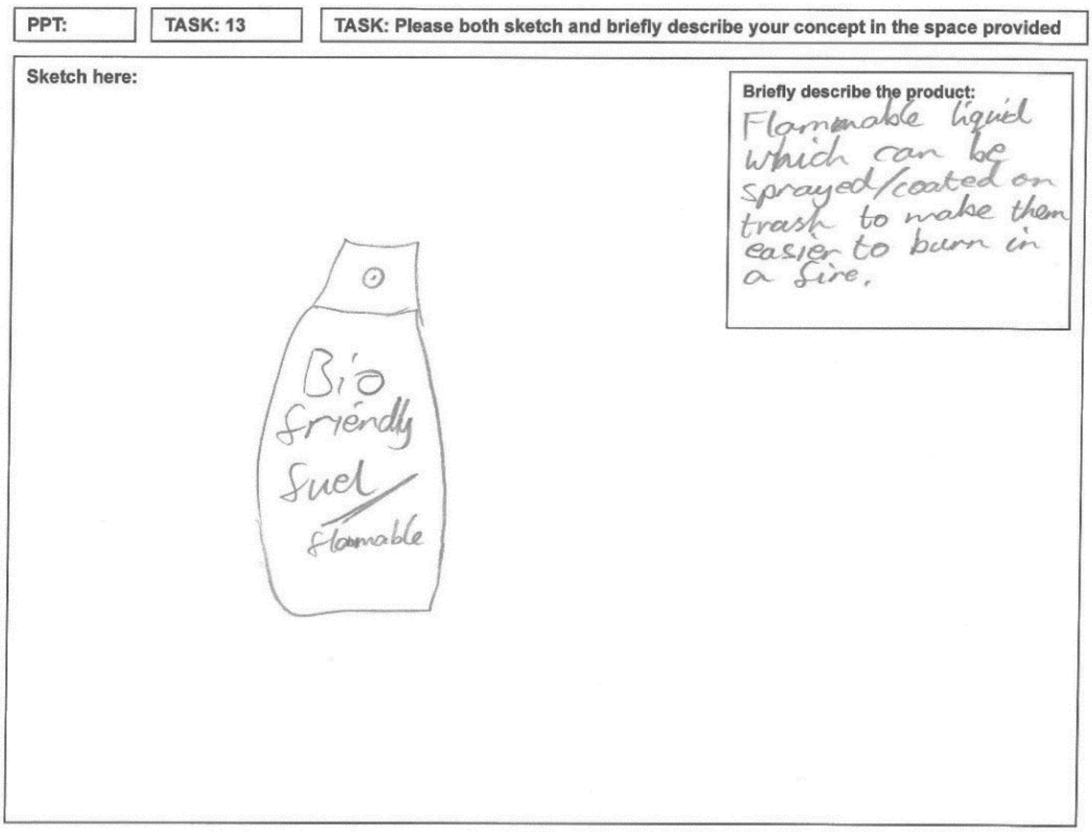


Figure 9-4 - An example of a non-combinational, ‘generation’ response. The description reads Flammable liquid which can be sprayed/coated on trash to make them easier to burn in a fire”.

Correlational analyses showed that as base-concept similarity increased, the proportion of non-combinational responses increased and the proportion of combinational responses decreased, whilst the proportion of ‘other’ responses was not significantly associated with similarity. The increase in non-combinational responses may have been caused by high similarity base concepts. The designers may have found it difficult to combine highly similar pairs of design concepts, either because they already had many features in common (thereby inhibiting featural combination) or because they found it difficult to find a plausible way to place them into a scenario (thereby inhibiting relational combinations). When faced with this difficulty, the designers may have changed their standards of success for the task to prioritise the successful creation of a design concept by any means over strict adherence to the instruction to ‘combine’ the base concepts. Other ideation processes like generation and transformation could act as a ‘release valve’ for producing a design concept when faced with highly similar pairs. This may occur when the designer finds it impossible to produce a combination, or when they feel that any combination they could produce would be a trivial change relative to the base concepts. It would not be surprising that designers who typically carry out concept ideation using a variety of cognitive processes might fall back on other processes to complete a task when faced with difficulties.

A counterpoint to the process-switching explanation is that non-combinational responses were not rated as being more difficult to produce (Section 8.1.4.3), although the 'Other' codes were associated with increased difficulty. However, since difficulty self-ratings were taken *after* the participants had combined each pair, they may have already alleviated any excess difficulty by switching to a non-combinational process, causing them to rate the trial as no more difficult than normal.

Despite their prevalence in the design concept combination task, no examples of non-combinational processing were found in the conceptual combination literature (Section 2.3) and none appear to have been reported in the other two design combination experiments that have measured combination type (Doboli *et al.*, 2014; Nagai *et al.*, 2009). If the high similarity is the cause of a process-switching effect, why has it not occurred previously? One explanation is that the most similar *design concepts* used in Combo-P3 may have been relatively more alike than any of the *category concepts* used in prior studies. In prior experiments, both in linguistic conceptual combination and design combination, participants were presented with pairs of object concepts like HAMMER-SAW or DESK-GUITAR. Unless identical or near-synonymous concepts were used (e.g., seat-chair, or drink-beverage), the use of two different words implies that the concepts represent two distinct categories of objects in the real world, placing an inherent limit on maximum similarity. In *design concept* combination, because the stimuli are not represented by single category labels, there is no inherent limit on maximum similarity. Highly similar design concepts can have many features that are typical of members of the same category of objects. As an example of the relatively high potential similarity of design concepts, consider the pair of base concepts shown in Figure 9-2. Both design concepts represent products for repurposing camping waste as fuel and could be classified as 'stoves'. The two concepts may have relatively high similarity by virtue of their common status as stoves: they both burn things, receive waste, have an internal storage volume, and expel heat. This relatively high degree of similarity would not have been found in prior studies that used category concepts represented by words as stimuli (Doboli *et al.*, 2014; Nagai *et al.*, 2009). This very high similarity may have made them difficult to combine, thus leading to the non-combinational responses.

Overall, it is plausible that the non-combinational responses were the result of the designers utilising other ideation processes. However, the alternative explanation for the non-combination responses is that combination *did* occur, but that it was not detectable by the independent raters and so the analyses missed a large proportion of 'genuine' combinational responses. For example, designers may have intentionally combined abstract features such as weight or cost in the process of making the apparently 'non-combinational' responses. This would mean that the decision to remove the non-combinational responses from the analyses has misrepresented the outputs of design concept combination and their relationship with similarity. Given the prevalence of these responses at nearly 20% of all successful forms of ideation, future research would benefit from including additional measures to probe the origins of non-combinational responses and confirm why they occur. Introspective methods may provide additional insights into the cognitive processes involved in combining stimuli.

9.2.1.2 The outputs of design concept combination

If design concept combination occurred via dual-processes of comparison & construction and scenario creation, it was expected that (after removing non-combinational responses) there would be three kinds of combined concept. Feature-mapping and hybrid combinations would stem from the comparison and construction process, and relational combinations would result from the scenario creation process. Table 9-3 compares the expected combination types with those identified during coding. In Combo-P3 three types of combination were identified. Relational combinations were identified as expected, and a general 'featural' type of combination was identified, but it was not possible to distinguish between the expected feature-mapping combinations (the transfer of a single feature) and hybrid combinations. An unexpected combination type was also identified; ambiguous combinations were so named because they could be construed as either featural or relational, depending on the interpreter's perspective.

Table 9-3 – A comparison of the expected combination types and those that were identified in Combo-P3

Cog. Pro.	Combination type	
	Expectations from the proposed model	Combination types from Combo-P3
Comparison & Construction	**Feature-mapping combinations** *a single feature is transferred from one design concept to the other*	**Featural combinations** *The newly created design artefact contains unique elements from each of the base concepts.*
	Hybrid combinations *a design concept is created that is a mixture or half-and-half of both base concepts*	
	N/A	**Ambiguous combinations** *Can be coded as featural or relational.*
Scenario creation	**Relation interpretations** *Design concepts are placed into a scenario, related by an external relation.*	**Relational combinations** *The combined concept has been created by placing an entity from one design concept into an external, complementary relation with an entity from another design concept.*

Note: Cog. Pro. = cognitive process.

The featural and relational combination types are consistent with the proposed cognitive processes. A comparison and construction process could give rise to the featural combinations, and a scenario creation process could give rise to the relational combination. However, the ambiguous combinations, as well as the lack of a distinction between two kinds of featural combinations, are inconsistent with the proposed model.

The inability to code hybrid combinations meant that the algorithmic-level predictions about the proposed Structural Alignment process could not be tested (Section 8.1.5.2). Hybrid combinations are a key part of Wisniewski's rationale for the existence of a hybrid process, and their absence in design concept combination reduces the strength of evidence for the proposed model.

It may be the case that 'hybrids' do not constitute a plausible type of combination for combining *design concepts.* A key property of the kinds of category concepts used in prior research is that words correspond directly to concepts (Ran and Duimering, 2009). This is important for conceptual combination because this correspondence is used in studies of conceptual combination when interpreting combination types. For example, consider a person who is asked to interpret the words 'robin snake'. Here, the word 'robin' corresponds to a person's psychological concept of a robin, and likewise for the word ‘snake’. They may respond with a featural combination such as “a snake with a red underbelly” (Wisniewski, 1997, p.169). Alternatively, they may interpret a ‘robin snake’ as ‘an animal that is a cross between the two - half robin and half snake’, thereby creating a hybrid combination (example modified from Wisniewski, 1997, p.169). In both examples, the changes that have been made to the base concepts by the interpreter are explicitly communicated in the response. In the featural combination, the feature ‘red’ has been attributed to the snake. In the hybrid example, the phrase ‘a cross between’ tells the researcher that a hybrid combination has occurred.

In *design concept* combination, the entities being combined are not category concepts, are not represented by one-word category labels, and (unless explicitly stated by the designer) it is often difficult to quantify the number of features that have been adopted from the other base concept. This in turn makes it difficult to distinguish between feature-transfer and hybrid combinations. For example, Figure 9-5 below shows a response to the base concepts in Figure 9-2. The combined concept is described as a “portable rubbish burner that changes heat from waste into usable heat”. Visual inspection of the sketch shows features of both base concepts; the cylindrical form of the left-side base concept, and the front-loading aspect of the right-side base concept. Beyond this, however, there is no language to indicate whether a single feature has been transferred between the design concepts, or the designer has created something that is a ‘cross between’ or a ‘hybrid’ of the two base concepts. Without the convenience of a short written description that refers to either of the two base concepts, it was not feasible to distinguish between two kinds of featural combination.

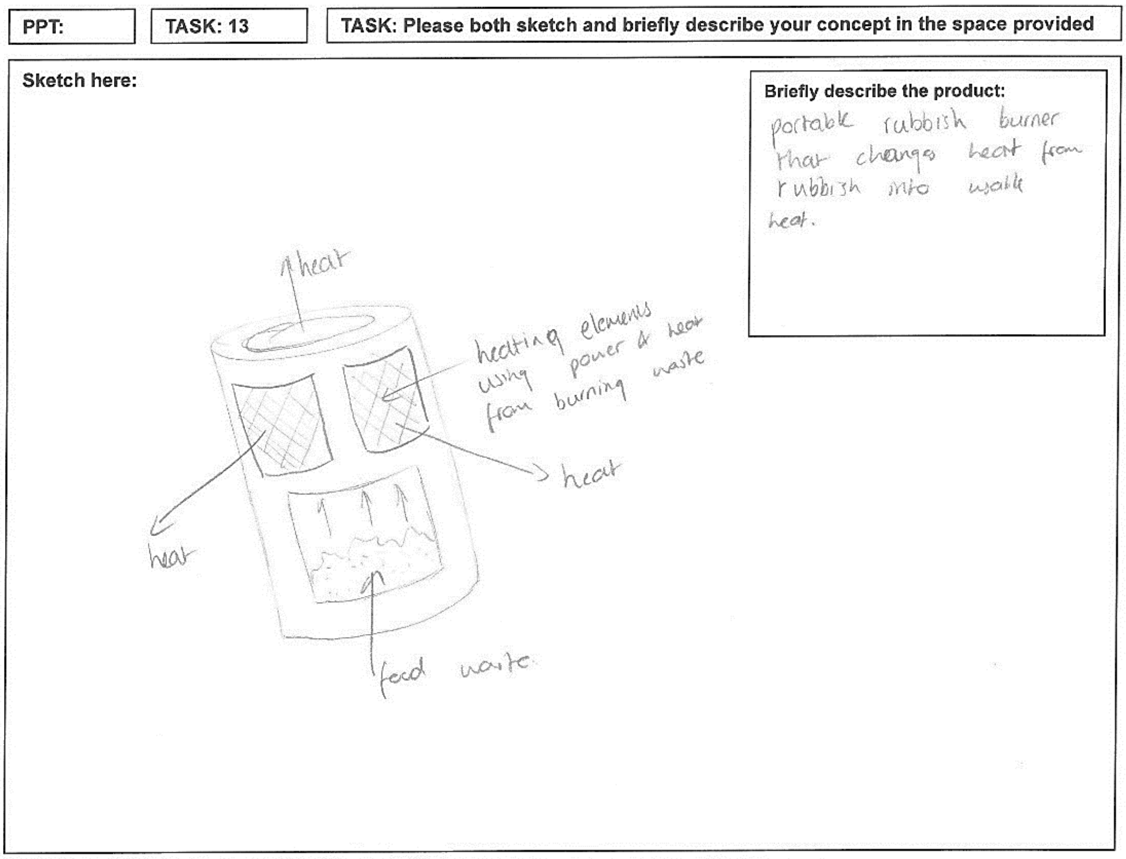


Figure 9-5 – A response to base concept pair Sc_13 (Figure 9-2)

The ambiguous combinations may also be caused by characteristics found in design concepts but not category concepts. That is, ambiguous combinations appear to arise when an entity in one base concept can fulfil the role of an entity in the other base concept; something that only happens when there are multiple entities in one or both of the base concepts. In the example of an ambiguous combination shown in Figure 8-4 (p.216), each of the base concepts Sc_25a and b (Figure 8-1, p.213) comprise scenarios of entities. Sc_25a is a 'BUS *carrying* CONTAINER' and Sc_25b is a 'DRONE *carrying* WATER BOTTLES'. The ambiguous combination can arise when an entity in one base concept can be 'swapped' for an entity in the other base concept that can fill the same role. In this case, the combined design concept is a 'DRONE *carrying* CONTAINER'. These kinds of combinations were not accounted for by the original model, because the conceptual combination model that it was based on was developed on word-word combinations that did not leave room for the combination of scenarios. The ambiguous combinations imply the need for a new cognitive model that accounts for the combination of pairs of scenarios.

Taken together, the presence of ambiguous combinations, the lack of hybrid combinations, and the unique characteristics of design concepts show that the proposed model is not adequate. However, the ambiguous combinations appear to reflect a process of entity substitution, where an entity in one base concept is swapped for an entity in another. A triadic model of 'featural', 'relational' and 'entity substitution' combinations could form the basis of a new model of design concept combination.

9.2.1.3 The relationship between similarity and combination type

If design concept combination occurred via two processes involving comparison and scenario creation, then it was proposed that the relationship between similarity and combination would mirror that of the Dual-Process model of conceptual combination. The relationship was represented in terms of four claims (Section 4.1.3) concerning the relationship between similarity and combination types. The analysis found no apparent relationship between similarity and the proportion of ambiguous combinations, but there were statistical and visual relationships between similarity and the proportion of featural and relational combinations.

The first claim made by the proposed Dual-Process model is that similarity facilitates featural combinations. According to Wisniewski's reasoning, as similarity increases (and by association so does alignability), it should be increasingly easier for people to find features to transfer between concepts. By this reasoning, it should be increasingly easy to produce featural combinations of design concepts. This was supported by the positive correlation between similarity and featural combinations. This is consistent with prior empirical findings from linguistic conceptual combination experiments. Using a dichotomy of high and low similarity stimuli, Wilkenfeld and Ward (2001) found that property and hybrid combinations (collectively analogous to featural combinations in this thesis) were more prevalent for high similarity pairs and 'relation linking' (analogous to relational combinations in this thesis) were more prevalent for low similarity pairs.

The second claim is that relational combinations occur in the presence of a plausible thematic relation. It was found that despite the statistically significant correlation between similarity and the proportion of relational combinations, the frequency of relational combinations varied substantially for pairs of design concepts with almost identical similarity ratings. This would be expected if the proportion of relational combinations was being driven by the ease with which designers can identify plausible thematic relations.

The third claim is that featural combinations occur in the absence of thematic relations. This is supported by the finding that all pairs of design concepts received at least one featural combination, whereas only 25/30 received at least one relational combination.

This would be expected if 5 pairs had no plausible thematic relations between them, causing the participants to resort to featural combinations instead.

The fourth claim is that relational combinations are inhibited by very high base-concept similarity. This is consistent with the negative correlation between similarity and the proportion of relational combinations. Additionally, visual inspection of Figure 8-16 showed a partial suppression of relational combinations for the most similar base concepts. Of the five pairs of base concepts that received no relational combinations, all were in the top third of rated similarity the top three received no relational combinations at all.

Taken together, it was possible to find evidence in support of all four claims. The relationship between similarity and the proportion of featural and relational combinations is consistent with empirical research in conceptual combination, the claims of Wisniewski's Dual-Process model, and thus the model that was proposed in Section 4.1.3. This suggests that there may be common effects associated with design concept combination as is found in linguistic conceptual combination. From the perspective of a Dual-Process model, this would imply a facilitative effect of alignability on featural combinations and the involvement of thematic relations in the production of relational combinations.

9.2.2 Consistency with prior knowledge of design combination

The results from this thesis contribute to knowledge about how designers combine ideas to create new ones. Knowledge has been gained for the first time about the outputs of *design concept* combination. Like prior studies, knowledge has been gained about the types of combinations that designers produce. Unlike prior studies, the research in this thesis is the first study to measure the relationship between similarity and combination type in design. Prior research that has coded combination types has not measured similarity (Doboli *et al.*, 2014; Nagai *et al.*, 2009) and prior research that has measured similarity has not coded combination types (Chan and Schunn, 2015; Jang, 2014; Jang *et al.*, 2019)[18]. Thus far, all studies that have coded combination types have been derived

[18] Doboli *et al.* (2014) claim that they manipulated the similarity of their stimuli, but they did not report any measurements of base-concept similarity or state how they operationalised similarity.

from Wisniewski's Dual-Process model to varying extents. This provides a datum for comparison to highlight how the research in this thesis compares to prior findings.

Table 9-4 compares the kinds of combinations coded in this thesis with those from prior works. Prior studies used category concepts, whereas the research in this thesis used *design concepts.* A new dimension that becomes important when attempting to compare these three studies is the relative level of system decomposition of the stimuli, i.e., whether the designers are combining parts or wholes. For example, electronic components (Doboli *et al.*, 2014) are the building blocks of other products, whereas the design concepts in this thesis and existing products (Nagai *et al.*, 2009) are themselves composed of smaller parts.

Table 9-4 – A comparison of combination types identified in design in three studies

		(Nagai *et al.*, 2009)	**(Doboli *et al.*, 2014)**	**This thesis**
Concepts	**Type of concepts**	Category concepts	Category concepts	Design concepts
	System level	Whole	Components	Whole
Combination types (using Wisniewski's terminology)	**Property**	Property-mapping *Transfer of a single feature*	Property-based combination (PBC) *Transfer of a single feature*	Featural combinations *Any mix of different features from both base concepts*
	Hybrid	Hybrids *Conjunctions comprising multiple features of both base concepts*	n/a	
	Relational	Thematic *Indirect combination via abstraction and instantiation*	Relation-based combination (RBC) *Direct external relations*	Relational combinations *Direct external relations*
	n/a	n/a	n/a	Ambiguous

The table shows the overlap between what Wisniewski termed 'property' and 'hybrid' combinations. The research in this thesis confirms previous findings that designers can create new design concepts that comprise the constituent features of existing products or prior designs. It is not surprising to find these kinds of combinations. It is the kind of combination prescribed in design methods such as morphological combination, where designers combine numerous combinations of sub-function solutions to create overall products (Zwicky, 1967). Examples can also be seen in protocol studies, where designers produce these kinds of combinations naturally during conceptual design (Section 2.2). These kinds of combinations predominate when combining 'whole' entities, but are infrequent when combining product components. Featural combinations occurred in 60.25% of all instances of combination in Combo-P3, which is consistent with the combination of category concepts where 75.61% of the combinations were property mapping or blending (Nagai *et al.*, 2009). In contrast, the proportion of PBCs (Doboli *et al.*, 2014) was low (0.04%).

A second kind of combination coded from design concept combination was termed the 'relational' combination, wherein the newly combined design concept comprised entities from different base concepts placed into scenarios. This kind of combination is derived from the 'relation-linking' combination in the Dual-Process model of conceptual combination (Wisniewski, 1997a). In design, two other kinds of combinations have been derived from Wisniewski's relation-linking process. They are the thematic combinations (Nagai *et al.*, 2009) and the relation-based combinations (RBCs) (Doboli *et al.*, 2014). Relational combinations occurred in 24.4% of combinations of category concepts (Nagai *et al.*, 2009) and 26.7% of design concept combinations (Combo-P3). In contrast, almost all (95.9%) combinations of electronic modules were coded as RBCs (Doboli *et al.*, 2014). A difference across these combination types is whether the external relation is direct or indirect. The relational combinations in this thesis and the RBCs (Doboli *et al.*, 2014) can both be described as *direct* thematic relations. Direct combinations involve relations between two base concepts. That is, base concepts are placed into scenarios or systems with e.g., spatial or functional relations. The thematic processing defined by Nagai *et al.* (2009) is *indirect* in that the base concepts are first used to create a scenario and then a new design concept is generated for use within that scenario. An example is given of the thematic combination of SNOW-TOMATO resulting in a 'refrigerator that humidifies the food in it' (Taura, 2016b). In this example, an individual first creates a scenario from the scene of a tomato stored in or preserved in snow. They may then extrapolate that the high moisture content of the snow may preserve the freshness of the tomato, leading to the idea of a refrigerator that preserves food by maintaining a humid atmosphere. The research in this thesis appears to be the first documented empirical evidence of combination via *direct* external relations between *design concepts.*

The contrasting proportions of combination types between the study of electronic component combination (Doboli *et al.*, 2014) and the others can be attributed to differences in the relative level of system decomposition of the stimuli used in the experiments. Featural combinations are the predominant form of combination when combining 'whole' entities but are both rare and low quality in the combination of product components. Relational combinations occur at relatively low frequencies when combining 'whole' concepts, but are the predominant form of combination for product components. One issue with this is that all three coding schemes were derived from the same theoretical model, but what is coded as a relational or thematic combination in some contexts may be coded as a featural combination in others, depending on what level of system decomposition the designer is thought to be starting at. This is problematic for two reasons. First, it limits the generalisability of concept-type coding schemes. Outside experimental conditions, it may not be so easy to define the initial conditions of an instance of combination, especially when designers can pick from a picture of product components and 'whole' sources of inspiration. Secondly, these codes are supposed to reflect two distinct cognitive processes, but if the only thing that differentiates them is the system-level of the initial stimuli, then it is not clear what meaningful distinction there is between featural and relational processing. Future research would benefit from proposing cognitive models of combination that account for combination at multiple levels of system decomposition.

Overall, design concept combination may be said to relate to other forms of design combination in four ways.

- Designers produce featural and relational combinations when combining category concepts and design concepts, but the definitions of these combination types differ depending on the level of system decomposition of the stimuli being combined.
- Featural combinations are the predominant form of combination when designers are combining 'whole' category concepts and design concepts, but are infrequent and low quality when combining product components.
- Relational combinations can occur via direct combination or by first creating an abstract scenario constituent concepts and then creating a new product within that scenario. When combining whole concepts, relational combinations are less frequent than combinations involving feature transfer or hybridisation, but they are the predominant form of combination when combining product components.
- Ambiguous combinations only occur when combining design concepts, and may be the result of a disparity between the number of entities between two base concepts.

9.2.3 Validity

Knowledge has been gained about the outputs of design concept combination and their relationship with similarity. There are three main limitations associated with this knowledge. First, the confirmationist approach to studying the computational level phenomena meant that a conscious effort was made to look for patterns in the data that supported the claims of the proposed model. This means that any support *for* the Dual-Process model may in turn be subject to biased interpretation. Secondly, much of Wisniewski's initial evidence for the model (Wisniewski, 1997a) and defence of the model (Wisniewski, 2000, 2001) relies on qualitative, verbal claims. This makes it easier to find patterns in the data (through confirmationist reasoning) that verify those claims. Third, the inability to test the planned hypotheses means that none of the work directly probes algorithmic-level processing.

Because Combo-P3 was correlational and the planned hypothesis tests associated with the algorithmic level could not be conducted, it is not possible to draw any inferences about causality between similarity and combination type. In the proposed model, the relationship between similarity and combination type is a product of four factors: a facilitating effect of alignability on featural combinations, the presence or absence of thematic relations, a default to featural combinations in the absence of thematic relations, and the inhibition of plausible thematic relations at high similarity. In this view of combination, similarity is a useful *indicator* of the underlying processes, but similarity is not the causal mechanism that determines combination type. Rather, the degree of alignability and the presence or absence of a plausible thematic relation are the determinants of combination type. However, these four factors are not necessarily unique to dual processes of comparison and scenario creation. For example, in conceptual combination, the ECCo model of conceptual combination (Lynott and Connell, 2010) also accounts for similarity with similar mechanisms.

The validity of inferences about combination processing is determined in part by the generic issues of construct validity discussed in Section 9.3.2. The findings are currently limited to PDE students in lab conditions combining pairs of early stage design concepts created in a concept generation experiment. Further research will be required to assess the external and ecological validity of the types of combinations that designers produce and their relationship with similarity.

Specific issues of generalisability relate to the measurement of combination and the inputs to combination. Measurement in the combination study is threatened by mono-operation and mono-method biases since only one measure of combination was used and only one task was used to study it. Construct validity could be improved by using multiple measures or tasks. An additional threat to construct validity was identified in the pre-analysis data checking (see the summary in Section 8.1.3.3). The proportion of combination types differed across the three design tasks, and so while the findings from Combo-P3 are valid for this sample of 30 pairs of design concepts, they do not necessarily generalise to all design tasks.

Construct validity is further limited in that the form of design concept combination studied in this thesis is an artificial and narrow example of all instances of design concept combination, and combination processing in design more generally. The framework of computational-level combination variables (Section 2.1.1) shows the variety of dimensions along which design concept combination can vary. Any variation along these dimensions may result in unexpected processing outcomes that have yet to be uncovered. For example, design combination is not necessarily limited to two input concepts. Designers may combine elements from multiple prior concepts as shown in Figure 2-4 (Gonçalves and Cash, 2021). Previous research indicates that research investigating combination types could be extended to the combination of multiple inputs (Doboli *et al.*, 2014), but the code definitions and coding procedure are not clearly reported and so it is difficult to gauge how well the concepts transfer to multiple stimuli. Another key limitation is that design concepts may vary in concreteness and detail. In later stages of the design process, designers may reason more about e.g., forces, materials and costs than they do when combining early-stage design concepts. Thus, the outputs of combination and their relationship with similarity may differ depending on the content of the knowledge being processed.

9.2.4 Implications

From the research in this thesis, we now know about the types of combinations that designers produce and how those types related to the relative similarity of the base concepts. This is basic, descriptive research, that is primarily intended to support the development of theory and act as a foundation for additional research into designer cognition. Unlike all prior research into design combination (Section 2.1.1.3), the findings do not measure any aspects of value or performance such as creativity, novelty or usefulness. Thus, it is not possible to make prescriptive claims for designers. Nonetheless, there may be benefits to educators and practitioners of knowing more about the fundamentals of combination processing in design.

One finding with implications for education is that design students in experimental settings frequently combine design concepts into scenarios (26.7% of the time). This finding may prompt educators to consider whether design education supports or constrains the natural behaviours of design students. For example, in the researcher's experience in the DMEM department at the University of Strathclyde, students in their design projects are tasked with designing single products rather than systems of connected artefacts. This is a reasonable constraint, given the limited time and resources available to any student project, students cannot be expected to prototype, embody and detail multiple products that work in a coordinated system. However, this focus on the creation of a single artefact may be unintentionally inhibiting the tendencies of novice designers to think about scenarios or systems of products. This might introduce biases into their behaviours as they develop through their careers, potentially leading them to avoid or undervalue system-based concepts.

Knowledge of the phenomena and effects of combination can be used in the development of computational ideation tools. Existing computational combination tools work by presenting designers with combinations of semantic categories represented by words or pictures. The Combinator combines category concepts by overlaying images or halving them and splicing the halves together (Han *et al.*, 2018b). The expert system developed by Luo *et al.* (2021) retrieves keywords at specified levels of semantic distance that the designer has to combine manually. Neither tool can provide designers with stimuli that are tailored for producing the featural or scenario combinations that designers produce in conceptual design. However, models like the Dual-Process model could provide the theoretical basis for new computational tools. For example, if the featural combinations identified in this thesis were produced by a comparison and construction process, then existing feature-based models of analogical reasoning such as DESSUA (Qian and Gero, 1992, 1996) could form the basis of automatic design concept combination tools.

9.2.5 Future research

To model how designers combine design concepts, future research would benefit from a focus on computational over algorithmic level knowledge, the development of formal models, improvements in research quality, the study of more naturalistic forms of combination, and the measurement of other outcome measures such as emergence and creativity.

Future research would benefit from establishing the phenomena of combination prior to the proposal and testing of algorithmic level models. As discussed in Section 8.1.5.2, the two hypotheses associated with the algorithmic level of the design concept combination model could not be tested because they were based on assumptions about computational level processing that were not borne out in the data. This demonstrates a lack of knowledge about the computational-level phenomena of design concept combination. To build a robust body of knowledge about computational level phenomena, the framework of combination variables (Figure 2-1) can be used to make researchers available of the variables associated with design combination and provide a common basis for situating the results of different studies.

An important development for outcome-based combination research will be to study combination in more ecologically valid settings. This includes combination involving more than two inputs, carried out by professionals during the course of the design process. Understanding how designers combine pairs of concepts can still be useful to aid the development of inspirational stimuli tools (e.g., Han *et al.*, (2018b); Luo *et al.*, (2021)), but it is not clear whether this generalises to other forms of combination (Section 9.2.3). Algorithmic-level models of design combination should be able to explain the phenomena associated with combining pairs of concepts as well as 3 or more concepts.

Beyond the measurement of combination types, other outcome measures such as emergence or creativity could provide a valuable contribution to understanding combination. At the computational level, knowledge of the factors that correlate with creativity are useful for developing prescriptive guidelines to aid students and professionals. A logical extension of the research in this thesis would be to explore whether featural or relational combinations differ in terms of creativity. This could be done through additional analyses or new experiments.

To develop algorithmic-level models of *design concept* combination, it will be beneficial for researchers to clearly establish their assumptions about how design concepts are represented in the mind. A key issue at the core of the stimuli creation process (Chapter 6) was a lack of knowledge about how design concepts are represented in the mind. Future research would benefit from deeper engagement with design artefact ontologies (Chakrabarti *et al.*, 2005; Rosenman and Gero, 1998; Štorga *et al.*, 2008; Umeda *et al.*, 1990). These ontologies provide a means of modelling the *representing world* in designer cognitive processing and generate hypotheses about designer cognitive processes.

9.3 General discussion of the thesis

This discussion provides a reflective discussion of the research in the thesis. This is structured around four topics: (i) the research findings and theoretical assumptions, (ii) the research approach, (iii) the research methods and (iv) the novelty and implications of the work. Each topic is discussed, where appropriate, in terms of the strengths and limitations of the research and what steps might be taken for future work.

9.3.1 Knowledge of the cognitive processes involved in design combination

Chapter 2 presented a review of literature on the cognitive processes involved in combination in design, creativity and conceptual combination. This provided an overview of the state of knowledge about combination cognitive processes at the outset of the research and led to the identification of gaps in knowledge and opportunities for new knowledge creation, thereby addressing Obj.1.

In comparison with other domains, namely creativity and linguistic conceptual combination, there was relatively little knowledge about how designers combine concepts to create new ones, especially at the algorithmic level of representation. As a starting point for research into combination processing in design, existing research was summarised into a framework of computational-level combination variables (Figure 2-1). The framework provides the first descriptive summary of the variables that have been investigated in empirical studies of combination in design. The framework benefits from including variables from multiple research methods (protocol studies and experimental paradigms). It can make researchers aware of the breadth and variety of variables that may influence their empirical studies of design combination. It is also likely that many of the variables associated with inspirational stimuli experiments (Vasconcelos and Crilly, 2016) would influence outcomes in combination and so both works could be integrated in the future.

A further output was the list of algorithmic-level processes implicated in combinational cognitive processing in design, conceptual combination and creativity (Table 2-8). This collection of models highlights the lack of diversity in algorithmic-level models in design. Research on combination in design predominantly draws from one of three sources:

Wisniewski's (1997b) Dual-Process model of conceptual combination (e.g., Doboli *et al.* (2014), Jang (2014), Jang *et al.* (2019), Nagai *et al.* (2009)); the research associated with the Creative Process Model (Mumford *et al.*, 2012) (e.g., Doboli *et al.* (2014)); or the Conceptual Blending framework (Fauconnier and Turner, 1998, 2003) (e.g., Nagai *et al.* (2009)). Yet, there is a breadth of theory that has been used to provide explanations for effects in other domains that do not appear to be considered in design. Future research in design could benefit from further integration with non-design domains to encourage more theory development.

The empirical research in this thesis has extended the previous state of knowledge by revealing the types of combinations that designers produce when combining design concepts and describing the relationship between similarity and output type.

9.3.2 Generic issues of validity

The experimental research in this thesis led to new knowledge about designer cognitive processes. The extent to which this knowledge generalises may be considered in terms of construct, external and ecological validity (defined in Section 4.2.5). Both the similarity and combination experiments share four study characteristics and thus share some threats to validity. The four constructs are participants, design concepts, IV manipulation and setting. The following discussion outlines the strengths and weaknesses associated with these four constructs and reasons about how external and construct validity might be affected in future research.

Participants. The experiments used students or recent graduates of product design engineering. Constraints placed on the minimum degree of educational experience (final semester of 2nd year or any year above) and the homogeneity of educational background (recruiting from a single degree stream) to improve internal validity. The downside of this homogeneity is that the students are not representative of PDE students globally due to their similar demographic and educational backgrounds, and are not representative of PDE professionals owing to their lack of experience.

External validity may be limited owing to individual differences in cognitive processes associated with similarity judgements. Simmons and Estes, (2008) have identified individual differences in similarity judgement processing and it has been shown that the relative preference for thematic and taxonomic processing changes across childhood (Smiley and Brown, 1979). Ecological validity may be limited owing to differences in deign cognition associated with design experience (Akin and Akin, 1996; Jansson and Smith, 1991; Lloyd and Scott, 1994) and domain of education (Kan and Gero, 2011; Purcell and Gero, 1996).

Design concepts. The use of design concepts as stimuli, rather than the kinds of category concepts used in prior studies, is a novel development in the study of designer combination cognitive processes. The design concepts used to create the pairs of base-concepts had two key strengths. First, drawing design concepts from more than one design task mitigated the threat to construct validity that design concepts generally were confounded with a single design task. This in turn improved external validity by making the results more robust to changes in design task. Secondly, since the design concepts were generated in a time-limited controlled experiment they were relatively homogeneous in their level of detail and development.

The limitations are that the design concepts are not representative of all forms of design concepts. This is both because they are early-stage design concepts and because they have been filtered to control for extraneous variables. It is not clear whether the results in any of the experiments would be replicable with physical prototypes of early-stage design concepts, functional models of complex systems, or CAD models of nearly-complete products. Nor is it clear whether the results would hold with design concepts that were randomly sampled from a set without filtering. Further, the trade-off to using design concepts created in lab settings it that they lack the ecological validity of design concepts generated during the design process. Similarly, the participants in the experiments did not create the stimuli themselves which would have been the case in a solo ideation session.

Stimuli similarity manipulation. The construct operationalised by the similarity manipulation was the *relative similarity* of pairs of design concepts. The similarity of the pairs was validated by human similarity judgements. Stimuli sets B and C, which were used in the main studies, were created to span as broad a range of similarity ratings as possible with two constraints; the most similar concepts should not be identical and the least similar concepts should, at minimum, address the same design brief.

The efforts to span the full range of relative similarity may have reduced construct validity by introducing implausible pairings. For example, the most similar concepts may be more similar than any designer would naturally select for combination during the design process. A further limitation is that the use of *pairs* of design concepts means that the findings do not necessarily generalise to the processing of more than two stimuli. It has been shown previously that designers may use elements of multiple prior concepts during combination (Chan and Schunn, 2015; Doboli *et al.*, 2014; Gonçalves and Cash, 2021).

Setting. Setting may be considered in terms of physical setting (the designer's location and the physical environment they interact with) and social setting (the presence or absence of other designers, stakeholders, clients etc.). The physical environment in which the experiment took place is representative of the spaces in which the participant sample typically conduct their design work, and is thought to approximate the kinds of environments in which professional designers work.

The presence of an experimenter has the potential to introduce biases such as the expectancy effect in which the experimenter subtly directs the participant through nonverbal communication, and the observer (Hawthorne) effect in which observation of the participant causes them to behave differently. The former was addressed using written task instructions to minimise the amount of interaction between the experimenter and participant. The latter was addressed by explicitly telling participants that they would not be observed unless otherwise stated in the task instructions and by the experimenter sitting as far away from, and at as much of an angle to the participants as possible. The social and design process settings do not reflect the social and design process settings in which students or professionals design. The methodology used by (Chan and Schunn, 2015) shows how similarity and combination can be measured from more naturalistic datasets.

9.3.3 Research approach

To address the aim of modelling the combinational cognitive processes of product design engineers in conceptual design, a hypothetico-deductive model-building research approach was adopted (Chapter 1). Two models were proposed that represented cognitive processing at the computational and algorithmic levels.

9.3.3.1 The hypothetico-deductive method and the nature of knowledge creation

The proposed models were developed through a process of analogical abduction. That is, the borrowing of explanatory principles from another domain in which a similar set of phenomena are better understood (paraphrased from Borsboom *et al.*, (2021)). A benefit of using existing models and their associated research methods is that the results of the studies of designer cognition could be compared to the findings from prior psychological studies.

Evidence for the models was gained through two kinds of knowledge creation. The first was through confirmationist reasoning; having proposed a model, exploratory quasi-experiments were conducted to look for evidence that verified elements of the model. Using this confirmationist approach provided some analytic freedom that would not have been available when conducting hypothesis tests. For example, it allowed for flexibility in coding the types of combinations that the designers produced without prematurely falsifying the proposed model. A limitation of this confirmationist reasoning is that the coding process is biased towards looking for pre-defined codes.

For each model, predictions were made based on the proposed algorithmic-level processes. If the answers to the research questions indicated some correspondence between the proposed models and the data, then further research would be conducted to test falsifiable hypotheses. Support for all of the hypotheses for a single model could be taken as support for the algorithmic-level process being tested, but falsification of any single hypothesis would negate support for that process. A benefit of this approach is that it enabled tests of cognitive processes that were assumed to be unavailable to introspection. By testing theoretically-derived hypotheses, it was possible to test for the existence of algorithmic-level processes that may otherwise have been unavailable to e.g., introspection.

In the case of the similarity model, the deductive approach and hypothesis tests were an appropriate way to investigate similarity cognitive processes (but see the future directions suggested in Section 9.1.5). The evidence from the explanation task (Sim-P2) that similarity judgements occur through feature-based processing provided a sufficient justification for moving on to hypothesis testing.

In the case of the combination model, the predictions made about algorithmic-level processing were premature. The results of the initial analysis showed that although the outputs of design concept combination were theoretically compatible with the proposed Dual-Process model, they were sufficiently different as to render the planned hypotheses tests redundant. This premature hypothesis testing reflects the proposal by (Scheel *et al.*, 2021), that "researchers who want to advance psychological science through hypothesis tests should spend less time testing hypotheses" (p.9). These authors propose that before making and testing hypotheses, it is beneficial to focus on defining concepts and the relationships between them, developing measures, outlining clear boundary conditions, stating auxiliary assumptions and specifying statistical predictions. More knowledge of computational level phenomena and effects is required before testable hypotheses about design concept combination can be generated.

9.3.3.2 Research scope

The research was motivated by the desire to understand designer combination cognitive processes. Extending the scope of the research to include similarity judgements had two key benefits. The first was that it was a reasonably efficient use of resources, given that there was a need to do some research on similarity anyway to create the stimuli. The second benefit was that by studying two cognitive processes with the same stimuli it was possible to make inferences about the existence of a stimulus compatibility effect by comparing across experiments (Section 9.1.1.1). A drawback to extending the scope was that the research programme became larger than was feasible given the intended length of the research programme. It may have been beneficial to focus on one cognitive process and approach the topic with more depth.

The study of *design concepts* was a response to an apparent gap in knowledge as existing lab-based research had only focused on the combination of category concepts. To the best of the researcher’s knowledge, the combination experiments in this thesis are the first to use *design concepts* as stimuli for controlled experiments studying combination. This is important because creative ideas are built on knowledge of existing category concepts but also previously generated ideas (i.e., design concepts). However, the study of design concepts introduced numerous challenges. No existing stimuli existed and so new stimuli had to be created which was time-consuming. Given the lack of knowledge about similarity judgements of design concepts, it was not clear what methods or models were the most appropriate for manipulating the similarity of design concepts. Further, the coding of differences and combination types required modifications of the coding schemes used in prior psychological studies of category concepts.

Given the null result of one of the hypotheses associated with the Structural Alignment model, it may also have been beneficial to first attempt to replicate the Structural Alignment model using the same stimuli as used in previous psychological studies (Markman and Gentner, 1993b, 1996) with PDE students before attempting to test the model with design concepts as stimuli. This would enable us to infer whether the null result is associated with design concepts in particular, or with some methodological aspect of the research.

9.3.4 Research methods

9.3.4.1 Literature review

Chapter 2 presented a review of literature on the cognitive processes involved in combination in design, creativity and conceptual combination. The literature review conducted in the thesis aligns with the generic definition of a 'literature review' given by (Grant and Booth, 2009). It was used to provide an examination of the state of knowledge at the outset of the research project. This directly led to the summary of combination cognitive process research across three domains, the identification of gaps in knowledge and opportunities for future research, the creation of the computational level framework of combination variables and the list of algorithmic level cognitive processes (Chapter 2) and the summary of the Dual-Process models of similarity judgments and conceptual combination (Chapter 3).

The main strength of the review is its breadth across domains (addressing three domains of research) and within each domain (focusing on a broad range of models of creativity and conceptual combination). This provides a more complete picture of the current state of knowledge about combination cognitive processes compared with prior research in design that all focus on a smaller number of models (Chan and Schunn, 2015; Doboli *et al.*, 2014; Jang, 2014; Nagai *et al.*, 2009).

The limitations of the review are that it is non-exhaustive and lacks systematicity. For example, it did not extend to cover the conceptual blending framework. Conceptual blending was omitted from the review in Chapter 2 because it does not make any testable predictions (Coulson and Oakley, 2001; Gibbs, 2001) and did not appear to offer a path toward model development in design. The original authors note that "conceptual blending is not a compositional algorithmic process and cannot be modelled as such even for the most rudimentary cases. Blends are not predictable solely from the structure of the inputs" (Fauconnier and Turner, 1998, p.136).

9.3.4.2 Stimuli creation and stimuli sets

Two methods were used to manipulate the similarity of pairs of design concepts. The first method was to code the design concepts and manipulate the number of common and different features. The second was to categorise the design concepts according to common purposes, functions and means. This second method was then edited and revaluated using human similarity judgements to create a third set of design concepts.

Stimuli Sets B and C (Sections 6.3 and 6.4) are sets of early stage, sketch- and text-based design concepts that vary in terms of their relative within-pair similarity. Both sets of design concepts may be used in future experimental studies. Set C may be considered a higher quality set as it has undergone an additional round of filtering to control for extraneous variables. The design concepts used to create each pair were created by designers from the same population, in a concept generation task with time constraints. The concepts are early-stage design concepts, represented by sketches with annotations and written descriptions. The strengths and limitations of the stimuli sets are generally the same as those discussed in Section 9.3.2 pertaining to the selection of design concepts and the similarity manipulation.

Overall, the key strength of the similarity creation methods was that they were theoretically grounded with respect to the representational assumptions in each specific phase. The drawbacks were that iterative testing of stimuli creation methods was time-consuming. The filtering of design concepts could also have been improved by using independent judges to check the application of the exclusion criteria in each phase.

Future research would benefit from the production of standardised banks of design concepts for use as stimuli in experiments; similar sets have been created for category concepts in psychology research (Battig and Montague, 1969; Brodeur *et al.*, 2010) In the future, a dedicated research programme could investigate additional methods of manipulating the similarity of pairs of design concepts. A more pragmatic approach would be to use intuitive manipulation of similarity and independent validation by human judgements. At present, too little is known about similarity judgements of design concepts to facilitate *a priori* control over alignability.

9.3.4.3 Correlational quasi-experiments

The main research method was the correlational quasi-experiment. A key benefit to the method is higher internal validity than non-experimental designs. The extensive filtering and iterative testing of the similarity manipulation would not have been possible if analysing a naturalistic dataset as has been done previously (Chan and Schunn, 2015).

The key limitation of this method is the increased difficulty with making causal claims. The correlational research in this thesis fails the requirements of precedence of cause. For precedence of cause to have applied, it would have been necessary to manipulate the alignability of the stimuli (i.e., the common and different features) and *then* measure the relationship between alignability and similarity. The stimuli used in empirical phases 2 and 3 were only validated by similarity ratings, so there was no precedence of cause. Non-correlational experimental methods could be used in similarity and conceptual combination research in the future.

9.3.4.4 Follow up interviews

Semi-structured interviews (Saunders *et al.*, 2019, p.437) were used to probe participants' experiences with study procedures in Sim-P2. A pre-defined list of eight questions was used. The answers to the procedural questions confirmed a number of issues that were raised voluntarily by participants before the interview or were obvious from observation.

One participant provided an unexpected level of introspective detail about their thought processes. They explained that when judging the similarity of pairs of design concepts, they used specific criteria to arrive at their ratings, such as user experience, where the product is used, cost and capacity (i.e., the capacity of water storage). This demonstrates that interviews specifically, or introspection generally, could have been used to provide insights into the cognitive processes involved in similarity judgements. Perhaps similar insights could be drawn about combination processing.

9.4 Chapter summary

This chapter has presented three discussions; two specific discussions about each of the models that were proposed and tested, and one general discussion of the research in the thesis as a whole. In Sections 9.1 and 9.2 the evidence for the proposed cognitive models of design concept similarity judgements and design concept combination were discussed. In Section 9.3, the research documented in this thesis has been discussed in terms of the strengths, limitations, and suggestions for future work, thereby satisfying Objective 4. These strengths, limitations and suggestions are summarised in Table 9-5.

Table 9-5 – A summary of the general discussion of the thesis

Topic	Strengths	Limitations	Future work
Research findings			
Framework of design combination variables and list of algorithmic-level models across domains			
	• Largest collection of computational-level experimental variables found in the literature to date.	• Computational level variables not yet mapped to theories.	• May be extended to include variables that influence ideation using inspirational stimuli.
Mental representation of design concepts			
	• Made the connection between algorithmic-level concept representation and models of design artefact representation • Demonstrated the applicability of three design concept representation schemes for various theoretical and methodological situations	• A limited breadth of the literature on design concept representation was considered • The extent to which the P-FBS paths are truly sufficient for capturing designer artefact knowledge in similarity judgements and combination is not yet known.	• A dedicated research programme is required on the role of design concept representation in algorithmic-level cognitive models.
Model validity			
Participants	• Participants had minimum degree of design experience (1.5 years in post-secondary education • Participants had homogeneous training (recruited from a single degree stream)	• Design students are not representative of PDE professionals • Demographic homogeneity limits generalisability • No analyses of individual differences conducted	• Studies of individual differences in cognitive processing • Investigate effect of expertise on experiment outcomes

Topic	Strengths	Limitations	Future work
Design concepts	• Using design concepts (rather than category concepts) is a novel step forward for design cognition research. • Using design concepts from multiple design tasks mitigated risk of confounding design concepts with a single design task. • Design concepts were relatively homogeneous in degree of development owing to being created in controlled conditions and being filtered with exclusion criteria	• Early-stage design concepts are not representative of all forms of design concepts. • Using design concepts created in experimental setting reduced ecological / construct validity. • Study participants did not create the design concepts themselves.	• Extend research (similarity and combination) to include different kinds of design concepts. • Use design concepts from naturalistic settings. • Use design concepts created by the same participants that do the similarity judgements or the combination.
Similarity manipulation	• *A priori* similarity validated by human similarity ratings • Design concepts span a broad range of similarity ratings	• Very high similarity concepts may be implausible candidates for making similarity judgements or combining in naturalistic settings • Study participants did not create the pairs themselves	• Extend research to include presentation of more than 2 stimuli • Extend research to include processing of more than 2 stimuli (group similarity and combination of multiple design concepts)
Setting	• The physical setting may be considered generally representative of the kind of environments that designers work in • Experimenter effects controlled for by standardised task instructions. • Observer (hawthorne) effect mitigated by telling supervised participants that they were not being observed unless clearly stated in the task instructions.	• The experimental conditions lacked the context that would otherwise be afforded by being in a collaborative or social environment, as well as by being part of an ongoing design process.	• Extend research to more naturalistic settings, including social or collaborative dynamics and the context of a real, ongoing design process

Topic	Strengths	Limitations	Future work
Similarity measurement	• Two measures of similarity avoids mono-operation bias • Priming issues are limited by eliciting similarity ratings before listing of commonalities and differences	• Mono-method bias	• Measure similarity with other tasks such as speeded difference or forced choice
Combination measurement	• Priming issues are avoided by eliciting similarity ratings from a different group than the combination participants.	• Mono-method bias • Mono-operation bias	• Extend research to include additional outcome measures, including emergence and creativity.
Comparison measurement	• Direct coding of differences from a commonality and difference listing task, superior to prior measures of alignable and nonalignable differences in the design literature	• Coding reliability may be improved with additional training	
Research approach			
The hypothetico-deductive methodology and the nature of knowledge creation			
HD methodology	• Led to the introduction of new theory and methods in a design context	• May have introduced bias, e.g., in coding	• More descriptive / exploratory research is required before testing hypotheses about combination processes in design • Research into similarity models may benefit from starting with replication studies and extending existing models one variable at a time • Atheoretical, inductive studies can lead to insights without *a priori* theoretical biases

Topic	Strengths	Limitations	Future work
Confirmationist reasoning	• Allowed for flexible adaptation of model and interpretation of results without premature falsification • Descriptive knowledge can be reanalysed with alternative theories	• Weak evidence for proposed model, alternative theoretical explanations available	
Falsificationist reasoning	• Facilitated the testing of causal cognitive mechanisms that may not be available to introspection (similarity)	• Data are not likely to be useful in other theoretical contexts • Hypotheses relating to the proposed combination model were made prematurely and turned out not to be applicable in design • Testing a similarity model without first conducting a replication limits the ability to make inferences about why the model was not supported	
Research scope			
	• The study of design concepts is a necessary development for understanding designer cognition. • Inclusion of research into designer similarity judgements was an efficient use of resources that needed to be developed anyway. • Studying two cognitive processes enabled some cross-model inferences.	• The research programme was too large for the time and resources available to the researcher. • The principle reason for including the similarity research was to facilitate the testing of H-C2. This turned out not to be possible because design concept combination was not adequately described by the proposed model.	• Future research should focus in more depth on a single cognitive process. Attempting to model two cognitive processes is excessive for a research programme driven primarily by a single researcher.

Topic	Strengths	Limitations	Future work
Dual-Process models of similarity and combination			
	• Studying similarity as an independent variable is useful regardless of the theoretical context	• Other models may be more applicable	• Other models may be used as theoretical bases for research in design • Atheoretical, inductive approaches could be useful in selecting models for testing in design or developing new models.
Research methods			
Literature review			
	• Includes multiple domains: design, creativity, conceptual combination • Compares across domains	• Did not include conceptual blending framework	• Consider implications of conceptual blending framework in design
Stimuli creation			
	• Multiple methods trialled. • Methods evaluated against success criteria with human similarity judgements. • Methods based on, and discussed with respect to, clear theoretical assumptions.	• Iterative testing of stimuli creation methods was time consuming. • Independent judges could have been used to check filtering and exclusion process.	• Intuitive manipulation of similarity and iterative evaluation with human judges is, at present, the most appropriate method for manipulating design concept similarity. • A dedicated research programme could investigate additional methods of stimuli manipulation. • Researchers would benefit from sets of standardised design concepts for use as experimental stimuli.

Topic	Strengths	Limitations	Future work
Correlational quasi-experiments			
	• Useful initial exploration of the relationship between similarity and various outcome measures.	• Limits ability to make causal claims	• If the stimuli could be created by manipulating alignability rather than similarity, the precedence of cause requirement of causality would not be broken and the ability to make causal claims would be improved.
Follow up interviews			
	• Aided in refinement of methods. • Provided self-reports about participants' cognitive processing.	• Method could have been used to greater effect. The unexpected degree of introspection could have provided valuable insights earlier in the research process.	• Interviews could be used to elicit introspection from participants to greater effect in the future.

10 Conclusions

The aim of the research in this thesis was to *model designer cognitive process(es) of combination in conceptual product design engineering* (Section 1.2). Although the focus of the research was on combination cognitive processes, it was proposed that knowledge about design concept similarity judgements would aid in generating knowledge about design concept combination. Thus, the research involved two parallel streams of research with the objectives of modelling the cognitive processes in design concept similarity judgements and design concept combination. A deductive research approach was used to propose and test two cognitive models. It was proposed that the combination of design concepts and similarity judgements of design concepts both involved the same dual processes of comparison and scenario creation, and so two cognitive models were tested experimentally, using the same stimuli but different participants.

This chapter provides a conclusion to the thesis by summarising the contributions to knowledge (Section 10.1), presenting the conclusions relating to the main knowledge contributions in detail (Sections 10.2 - 10.4), and summarising the strengths, limitations, and recommendations for future work (Section 10.5).

10.1 Summary of contributions to knowledge

Six contributions to knowledge were gained from the research (K1 to K6, below). The contributions stem from reviews of the literature, the development of the experimental methods, or the results of empirical research.

There are three primary contributions of the research (K1-3).

K1. A descriptive framework of variables associated with design combination. Figure 2-1 shows the variables associated with the goals, inputs and outputs of design combination. These variables were extracted from protocol studies of conceptual design and (quasi-)experiments of design combination. Combination cognitive processes can be described in terms of 27 variables: 4 associated with the goals of combination, 11 with the inputs and 12 with the outputs.

K2. A cognitive model of design concept similarity judgements. In experimental settings, student product design engineers make similarity judgements of pairs of early-stage, sketch-based design concepts via a process of comparison. Consistent with psychological models of comparison-based similarity judgements, the similarity of a pair of design concepts is a function of the commonalities and differences of that pair, and as the relative similarity of a pair of design concepts increases, so too does the number of commonalities that designers can list for that pair. There is no evidence that designers make similarity judgements based on the presence of thematic relations between design concepts, but thematic judgements may sometimes intrude on comparison-based processing. A structural alignment algorithm was proposed to explain the comparison-based process, but the model was falsified and so no conclusions can be drawn about algorithmic-level processing.

K3. Knowledge of computational-level combination processing: the inputs, outputs and effects of design concept combination. When student product design engineers are tasked with combining pairs of early-stage, sketch-based design concepts to create new ones, they predominantly respond by combining the concepts, but also produce non-combinational responses by e.g., generating new ideas, using commonalities from both base concepts to create new concepts, or transforming one of the base concepts. When they *do* combine, their combinations can be classed as one of three kinds: feature-mapping, relational or ambiguous. In general, as the relative similarity of a pair of design concepts increases, designers are more likely to produce featural combinations and less likely to produce relational combinations (and vice versa). This general trend may be further characterised by three factors: (i) base concepts of near-identical similarity can vary substantially in the proportion of featural and relational combinations, despite the group-level trends, and (ii) there are few or no relational combinations at very high levels of base-concept similarity and (iii) all base concepts can be processed by featural combinations, but not necessarily by relational combinations. Overall, the outputs of design concept combination and the relationship between similarity and combination type are partly consistent with the Dual-Process model of conceptual combination (Wisniewski, 1997a) and with category concept combination in a design context (Nagai *et al.*, 2009). However, the ambiguous combinations produced during design concept combination are not captured by any prior model, suggesting that there is something unique about design concepts that is not captured by existing models.

K3 directly satisfies the aim of the research. Knowledge was gained about computational-level combination processing (the goals, inputs and outputs of combination) that contributes to a cognitive model of combination processing in design.

The three secondary contributions to knowledge are:

K4. A summary of algorithmic-level mechanisms in combinational processes in design, creativity and conceptual combination. Table 2-7 and Table 2-8 list algorithmic-level cognitive processes that have been implicated in combinational cognitive processes in design, creativity or conceptual combination. This is the first cross-domain summary of algorithmic-level models associated with combination. The algorithmic-level list can provide a basis for future theory development and interdisciplinary research into combination processes.

K5. Methods for manipulating the similarity of pairs of design concepts. A methodological contribution was made in the form of the work carried out to manipulate, *a priori,* the relative similarity of pairs of design concepts for use as experimental stimuli (Chapter 6). This contribution comprises a literature summary of methods for manipulating similarity and a transparent account of the implementation of two methods. The two methods that were trialled were (i) the manual manipulation of the common and different features of pairs of design concepts, and (ii) the creation of pairs based on ontological distance in categories organised around the purpose and function of the design concepts. Both methods provided some control over subsequent similarity ratings. The transparent reporting outlines the theoretical assumptions, requirements for success, the implementation of each method, an evaluation of how well those methods met the requirements, and a discussion of the strengths and limitations of each method. This can help other researchers to select methods for stimuli creation and avoid some of the pitfalls that were encountered in this thesis.

K6. Stimuli sets. Stimuli Set C (Section 6.4) comprises 30 pairs of design concepts, with 10 pairs from each of three design tasks. The pairs are of varying similarity both within each design task across the whole set. The similarity of the pairs was rated by product design engineering students and graduates. It is the first set of early-stage conceptual design concepts used as stimuli in experimental design research. It can be used in future research investigating similarity effects in combination.

The six contributions correspond to the four research objectives as shown in Figure 10-1. The three primary contributions (K1-3) are presented in greater detail in Sections 10.2 to 10.4.

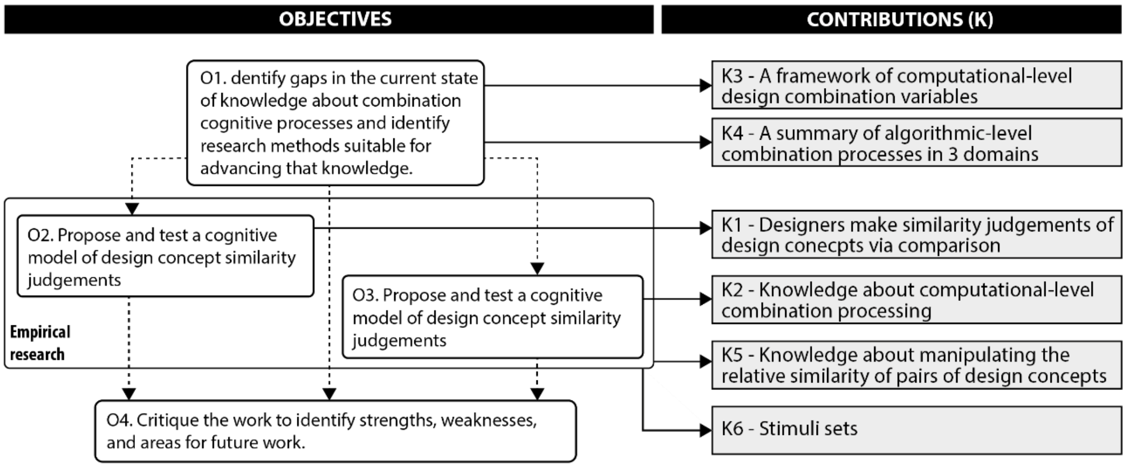


Figure 10-1 – Correspondence between the research objectives and knowledge contributions

10.2 K1 - A framework of computational-level design combination variables

A literature review of combination cognitive processes was conducted to satisfy O1, *identify gaps in the current state of knowledge about combination cognitive processes and identify research methods suitable for advancing that knowledge.* Combination was defined as the creation of design concepts based on two or more previously externalised entities. Combination was distinguished from generation, transformation and reasoning processes (Section 1.1.4.1). The review targeted design (O1.1) and non-design domains (O1.2), which were compared (O1.3) in terms of the computational and algorithmic level knowledge and the research methods used to create that knowledge.

The main knowledge contribution from this review was the descriptive framework of combination variables (Section 2.1.1). Combination cognitive processing in design can be described in terms of the goals, inputs and outputs of an instance of combination.

- The goal of combination is to produce new design concepts based on previously externalised entities. The goal is influenced by the designers' understanding of task instructions, their knowledge of design requirements and constraints, and performance goals (such as producing novel or useful outputs).
- The inputs to design combination can be described in terms of three classes of variables. Five variables relating to the 'type' of concept (*quantity, part/whole, internal world /external world, representation modality, taxonomic category*), three describing the properties of individual concepts (*ambiguity, associative effectiveness, visual complexity*), and three describing the properties of multiple concepts (*relatedness, similarity/distance, number of commonalities or differences*).
- The outputs of combination can be described in terms of three classes of concept. Seven variables describe the properties of individual concepts (*abstractness, overall creativity, novelty, practicality, quality, resistance to premature closure, usefulness*), one variable describes the properties of multiple outputs (*fluency*), and four variables describe how the outputs have changed relative to the inputs (*combination type, degree of reuse, elaboration, emergent features*).

The framework is the first integrative summary of empirical research into designer combination processes. It can be used in future empirical research into design combination to inform researchers about the variables that they may wish to manipulate or control.

10.3 K2 - The cognitive processes of design concept similarity judgements

Objective 2 was to propose and test a cognitive model of concept similarity judgements. A Dual-Process model of design concept similarity judgements was proposed 4.1) and a quasi-experimental research design was devised to test the model. The dual-processes were analogous to those described by the dual-process view of similarity judgements (Chen *et al.*, 2014; Estes, 2003a, 2003b; Estes *et al.*, 2011; Wisniewski, 1997a). This model comprises the Structural Alignment model of comparison-based similarity judgements and a second scenario creation process. Comparison-based similarity is a product of the common and different features of a pair of concepts. Thematic similarity is a product of the perception that a pair of concepts can 'go together' and occupy complementary roles in an external relation.

The results of two experiments (Sim-P2 and Sim-P3) indicate that design students make similarity judgements about pairs of early-stage, sketch-based design concepts via a process of comparison, where the similarity of a pair of design concepts is a product of the common and different features of the concepts. There is no evidence to support the existence of a scenario-creation process. The proposed Structural Alignment algorithm was not supported.

The results of Sim-P2 (n=11) indicated that designers make similarity judgements based on the common and different features of a pair of design concepts. Self-reported explanations for numerical similarity ratings referred to the number of common and different features of a pair of design concepts, but none contained reference to a thematic relation. Unexpected instances of thematic processing were identified in the commonality and difference listing task. This is consistent with the view that thematic processing can intrude on comparison processing in some circumstances.

In Sim-P3 (n=35), five hypotheses were stated based on the predictions of the Structural Alignment model of similarity judgements (Gentner and Markman, 1994, 1997; Markman and Gentner, 1993b, 1993a). Four of the five hypotheses were supported, but one was not.

- In support of H1a, similarity can be predicted by the number of commonalities and differences listed for a pair.
- In support of H1b, alignable differences count more against similarity than nonalignable differences.
- In support of H2, as the similarity of a pair of design concepts increases, so too does the number of commonalities of that pair.
- In support of H4, alignable differences are more numerous than nonalignable differences.

The fifth prediction was not supported. It was expected (H3) that there should be a positive correlation between the number of commonalities of a pair of design concepts and the number of alignable differences. No significant correlation was found. Consistent with this finding, it should also have been the case that similarity was positively correlated with the number of alignable differences. Rather, there was a statistically significant negative correlation between similarity and the number of listed ADs.

Overall, the support for H1a and H2 is consistent with comparison-based models of similarity judgements, including the Contrast model (Tversky, 1977) and the Structural Alignment model. The support for these hypotheses, combined with the feature-based explanations identified in Sim-P2, indicates that designers make similarity judgements of design concepts via comparison.

Three explanations were provided for the lack of support for H2. Firstly, insufficient statistical power is possible, but the sample size was sufficient to have a 93% chance of detecting an effect the size of that reported in prior analyses (Markman and Gentner, 1996). Secondly, measurement noise may have masked the expected relationship via commonalities and alignable differences. Third, some latent phenomena may be making the expected relationship between Com and AD by inflating the number of ADs at low similarity or limiting the number of ADs at high similarity. Hypotheses and analyses to test this possibility were proposed.

10.4 K3 - The cognitive processes of design concept combination

Objective 3 was to propose and test a cognitive model of design concept combination. A Dual-Process model of design concept combination was proposed (Section 4.1.3) and a quasi-experimental research design was devised to test the model. The dual processes were analogous to those described by the Dual-Process model of conceptual combination (Wisniewski, 1997a). It was proposed that when designers combine design concepts, they do so through one of two processes. A comparison and construction process based on a structural alignment algorithm would lead to the production of feature-mapping combinations (where one feature is mapped and transferred to the other concept) or hybrid combinations (where the resulting design concept is a mixture of both base concepts). A second process of scenario creation would produce relational combinations, where the two base concepts are combined via a complementary, external relationship.

The results of one experiment (Combo-P3) show that the design concept combination carried out by the student participants is partly consistent with the proposed model. However, there is sufficient divergence between the expectations of the model and the data to conclude that there are unique characteristics of design concept combination that are not captured by the proposed model. The knowledge contributions concern the inputs, outputs and effects of combination.

- Inputs and outputs: When designers are tasked with combining design concepts, they predominantly carry out combination as expected (70.5%), but frequently produce new design concepts that contain no evidence of combination (19.6%). When designers *do* combine, they produce three kinds of combinations. Featural combinations comprise features intrinsic to both base concepts. Relational combinations involve the placement of two entities from the initial base concepts into scenarios. Ambiguous combinations can be classified as either featural or relational combinations depending on one's frame of reference.

- Effects of combination: Correlational analyses showed that the proportion of featural combinations increases with base-concept similarity and the proportion of relational combinations decreases (and vice versa). Visual inspection of the data revealed three additional characteristics of the data. First, despite the correlational relationships, pairs of proximately similar design concepts can elicit contrasting proportions of featural and relational combinations. This suggests that there is a stimulus compatibility effect that influences the proportion of combination types for a given pair of design concepts. Second, there are few or no relational combinations at high similarity, indicating that very high similarity pairs inhibit relational combinations. Third, all base concepts received at least one featural combination, but not all received a relational combination, suggesting that designers produce featural combinations in the absence of plausible external relations.

The featural and relational combinations and their relationship with similarity are consistent with the proposed Dual-Process model. However, ambiguous combinations are not accounted for by a dual-process view of combination, and it was not possible to distinguish between two kinds of featural combinations (i.e., feature-transfer and hybrids). This means that there is insufficient evidence to conclude that design concept combination occurs via the proposed dual processes. Additionally, a consequence of not being able to code hybrid combinations is that the planned hypothesis tests associated with algorithmic-level processing could not be tested. Thus, it is not possible to make claims about how designers produce featural, relational or ambiguous combinations.

With regards to the original proposal that similarity judgements and combination share the same dual processes, similarity and combination *could* share a common cognitive basis, but combination likely involved additional cognitive processes. Similarity judgements appear to be carried out via comparison, and featural combinations could plausibly be a product of a comparison-based process. However, the relational and ambiguous combinations are not accounted for by a comparison model, and so it can be concluded that for design concepts, similarity and combination do not share the same underlying dual processes.

At the algorithmic level, it is unclear exactly how designers make comparison-based similarity judgements or produce featural combinations. The falsification of the Structural Alignment model of similarity judgements casts doubts over whether domain-general cognitive models provide ready-made representations of designers' cognitive processes of design concepts.

10.5 Strengths, limitations, and future work

Objective 4 was to critique research to identify strengths, weaknesses and areas for future work. This objective was addressed in the general discussion (Section 9.3) and summarised in Table 9-5. The key strengths and weaknesses of the research pertain to knowledge claims K2 and K3.

Knowledge claim K2 is that design students make similarity judgements of early-stage, sketch-based design concepts via a process of comparison. The strength of this claim stems from the multiple sources of evidence, namely the similarity explanations (sim-P2) and the correspondences between similarity ratings and the commonality and difference listing task (Sim-P3). Additionally, by drawing from the results of the combination experiments it was possible to rule out stimulus compatibility confounds, thereby strengthening the knowledge claim. Methodologically, the research that supports this claim is transparent and should be replicable. The manipulation of the stimuli was operationalised using human judgements of similarity and coding schemes were developed abductively to facilitate maximum alignment between the proposed schemes and the data gathered. The key limitations of K2 relate to the falsification of the Structural Alignment model, and the inability to draw conclusions about algorithmic level processing. Although three explanations were provided for the lack of support for the Structural Alignment model (statistical power, measurement noise or hidden confounds) it is not possible to decide amongst these explanations with the current data. One reason for this, other than the need for follow-up research, is the inability to rule out hidden moderators introduced from the few methodological changes introduced vs. the original research (Markman and Gentner, 1993b, 1996). That is, the use of typed rather than spoken responses, designers rather than non-designers, and design concepts rather than the kinds of stimuli used in previous studies. In the future, a more robust test of the Structural Alignment model could be conducted using a replication and extension approach. That is, by replicating the original studies by using the exact same materials and population, and then extending the research to include first designers and then design concepts.

The knowledge claim K3 comprises a series of specific claims about the inputs, outputs and effects of combination. This is the first ever contribution to knowledge about the combination of *design concepts* rather than e.g., category concepts or geometric stimuli. The strengths of the research relating to K3 are the transparency and depth of analysis that were conducted. The coding process was reported, starting with the definitions of combination types from the Dual-Process model of conceptual combination (Wisniewski, 1997a) and evolving iteratively to a coding scheme that is both consistent with the original model and fits the data from the design combination experiment Combo-P3. In the analysis of the relationship between similarity and combination type, base-concept similarity is transparently operationalised (Chapter 6). Further, the results of correlational analyses were not taken at face value, and visual inspection was used to more accurately characterise the relationship between similarity and combination type. The inclusion of a difficulty check means that it was possible to rule out a difficulty as a confounding influence on combination type. The main limitations of K3 stem from the deductive research approach. The use of an abductive coding approach may have introduced bias in the coding of combination types. Although the combination types could be coded reliably, alternative coding schemes that were not based on the Dual-Process model of conceptual combination could be more applicable. There may also be other variables that influence combination type such as diagnosticity that were not investigated in this research as it is not clear how this variable should be conceptualised for design concepts. Future research into combination cognitive processes may benefit from an inductive approach, where the outputs of combination are explored without any pre-existing coding schemes to better explore the landscape of combination types without any biases. An alternative path for research into cognitive processes would be to use measures of creativity (Baughman and Mumford, 1995; Mobley *et al.*, 1992; Mumford *et al.*, 1997; Scott *et al.*, 2005) rather than combination type.

In general, the strengths of the research were the interdisciplinarity, the novelty of studying design concepts and the attempts to model algorithmic-level cognitive processing via falsificationist reasoning. A key weakness of the work as a whole was its scope; the study of both combination and similarity judgements resulted in a research programme that was larger than suitable for the available resources and many open questions associated with each model that could have been answered with a more focused investigation of a single process. In the future, design researchers may be able to make use of hypothetic-deductive methodologies to investigate designer cognitive processes.

REFERENCES

Aamodt, A. and Plaza, E. (1994), "Case-based reasoning : Foundational issues, methodological variations, and system approaches", *AI Communications*, Vol. 7 No. 1, pp. 39–59.

Agresti, A. (2002), *Categorical Data Analysis*, 2nd ed., Vol. 482, John Wiley & Sons, Inc., Hoboken, NJ, USA, available at:https://doi.org/10.1002/0471249688.

Agresti, A. and Franklin, C. (2014), *Statistics: The Art of Learning from Data*, Essex, UK: Pearson.

Ahuja, G. and Lampert, C.M. (2001), "Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions", *Strategic Management Journal*, Vol. 22 No. 6–7, pp. 521–543.

Akin, Ö. and Akin, C. (1996), "Frames of reference in architectural design: analysing the hyperacclamation (A-h-a-!)", *Design Studies*, Vol. 17 No. 4, pp. 341–361.

Andrade, C. (2018), "Internal, External, and Ecological Validity in Research Design, Conduct, and Evaluation", *Indian Journal of Psychological Medicine*, Indian Psychiatric Society South Zonal Branch, Vol. 40 No. 5, p. 498.

Andreasen, M.M., Hansen, C.T. and Cash, P. (2015), *Conceptual Design: Interpretations, Mindset and Models*, *Conceptual Design: Interpretations, Mindset and Models*, Springer International Publishing, Cham, available at:https://doi.org/10.1007/978-3-319-19839-2.

Baddeley, A. (2012), "Working Memory: Theories, Models, and Controversies", *Annual Review of Psychology*, Vol. 63 No. 1, pp. 1–29.

Baddeley, A.D., Eysenck, M.W., Anderson, M. and Anderson, M.C. (2009), *Memory*, Psychology Press, available at: https://books.google.co.uk/books?id=3x-BPQAACAAJ.

Basadur, M. (1997), "Organizational Development Interventions for Enhancing Creativity in the Workplace", Vol. 31 No. 1, pp. 59–72.

Bassok, M. and Medin, D.L. (1997), “Birds of a Feather Flock Together: Similarity Judgments with Semantically Rich Stimuli”, *Journal of Memory and Language*, Vol. 36 No. 36, pp. 311–336.

Battig, W.F. and Montague, W.E. (1969), “Category norms of verbal items in 56 categories A replication and extension of the Connecticut category norms”, *Journal of Experimental Psychology*, Vol. 80 No. 3 PART 2, pp. 1–46.

Baughman, W.A. and Mumford, M.D. (1995), “Process-analytic models of creative capacities: operations influencing the combination-and-reorganization process”, *Creativity Research Journal*, Vol. 8 No. 776107160, pp. 37–41.

Benami, O. and Jin, Y. (2002), “Creative Stimulation in Conceptual Design”, *ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 251–263.

Bhaskar, R. (1975), *A Realist Theory of Science*, Routledge.

Biskjaer, M.M., Dalsgaard, P. and Halskov, K. (2017), “Understanding creativity methods in design”, *DIS 2017 - Proceedings of the 2017 ACM Conference on Designing Interactive Systems*, pp. 839–851.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), “Latent dirichlet allocation”, *The Journal of Machine Learning Research*, JMLR. org, Vol. 3, pp. 993–1022.

Blessing, L.T.M. and Chakrabarti, A. (2009), *DRM, a Design Research Methodology*, *DRM, a Design Research Methodology*, Springer London, available at:https://doi.org/10.1007/978-1-84882-587-1.

Boeijen, A. van, Daalhuizen, J., Zijlstra, J., van der Schoor, R. and Technische Universiteit Delft. (2010), *Delft Design Guide : Design Methods*.

Borsboom, D., van der Maas, H.L.J., Dalege, J., Kievit, R.A. and Haig, B.D. (2021), “Theory Construction Methodology: A Practical Framework for Building Theories in Psychology”, *Perspectives on Psychological Science*, Vol. 16 No. 4, pp. 756–766.

Brodeur, M.B., Dionne-Dostie, E., Montreuil, T. and Lepage, M. (2010), “The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research”, edited by Op de Beeck, H.P.*PLoS ONE*, Vol. 5 No. 5, p. e10773.

Brown, D.C. (2016), *Observations and Conjectures about Novelty Calculations*, *Design Creativity Workshop 2016*.

Bush, R.R. and Mosteller, F. (1951), “A model for stimulus generalization and discrimination.”, *Psychological Review*, American Psychological Association, Vol. 58 No. 6, p. 413.

Cacioppo, J.T. and Petty, R.E. (1982), “The need for cognition.”, *Journal of Personality and Social Psychology*, American Psychological Association, Vol. 42 No. 1, p. 116.

Campbell, D.T. (1960), “Blind variation and selective retentions in creative thought as in other knowledge processes.”, *Psychological Review*, American Psychological Association, Vol. 67 No. 6, p. 380.

Cash, P. and Kreye, M.E. (2017), “Uncertainty Driven Action (UDA) model: A foundation for unifying perspectives on design activity”, *Design Science*, No. 2003, available at:https://doi.org/10.1017/dsj.2017.28.

Chakrabarti, A., SARKAR, P., LEELAVATHAMMA, B. and Nataraju, B.S.S. (2005), “A functional representation for aiding biomimetic and artificial inspiration of new ideas”, *Ai Edam*, Vol. 19 No. 02, pp. 113–132.

Chan, C.-S. (1990), “Cognitive processes in architectural design problem solving”, pp. 60–80.

Chan, C.S.J. (2014), *The Impact of Sources of Inspiration on the Genesis of Creative Ideas*, University of Pittsburgh.

Chan, J., Fu, K., Schunn, C., Cagan, J., Wood, K. and Kotovsky, K. (2011), “On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples”, *Journal of Mechanical Design, Transactions of the ASME*, Vol. 133 No. 8, p. 081004.

Chan, J. and Schunn, C.D. (2015), “The importance of iteration in creative conceptual combination”, *Cognition*, Elsevier B.V., Vol. 145, pp. 104–115.

Chaudhari, A.M., Bilionis, I. and Panchal, J.H. (2019), “Similarity in engineering design: A knowledge-based approach”, *Proceedings of the ASME Design Engineering Technical Conference*, Vol. 7, available at:https://doi.org/10.1115/DETC2019-98272.

Chemi, T., Borup Jensen, J. and Hersted, L. (2015), *Behind the Scenes of Artistic Creativity: Processes of Learning, Creating and Organising*, Peter Lang International Academic Publishers.

Chen, Q., Ye, C., Liang, X., Cao, B., Lei, Y. and Li, H. (2014), “Automatic processing of taxonomic and thematic relations in semantic priming - Differentiation by early N400 and late frontal negativity”, *Neuropsychologia*, Elsevier, Vol. 64, pp. 54–62.

Christensen, B.T. and Schunn, C.D. (2007), “The relationship of analogical distance to analogical function and preinventive structure: the case of engineering design”, *Memory & Cognition*, Vol. 35 No. 1, pp. 29–38.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Academic press.

Coley, F., Houseman, O. and Roy, R. (2007), “An introduction to capturing and understanding the cognitive behaviour of design engineers”, *Journal of Engineering Design*, Vol. 18 No. 4, pp. 311–325.

Costello, F.J. and Keane, M.T. (2000), “Efficient creativity: constraint-guided conceptual combination”, *Cognitive Science*, Vol. 24 No. 2, pp. 299–349.

Costello, F.J. and Keane, M.T. (2001), “Testing two theories of conceptual combination: alignment versus diagnosticity in the comprehension and production of combined concepts.”, *Journal of Experimental Psychology. Learning, Memory, and Cognition*, Vol. 27 No. 1, pp. 255–271.

Coulson, S. and Oakley, T. (2001), “Blending Basics”, *Cognitive Linguistics*, Vol. 11 No. 3–4, pp. 175–196.

Cross, N. (1990), “The nature and nurture of design ability”, *Design Studies*, Vol. 11 No. 3, pp. 127–140.

Cross, N. (2001), “Design cognition: results from protocol and other empirical studies of design activity”, Elsevier, pp. 79–103.

Crossan, F. (2015), “Research philosophy : towards an understanding”, Vol. 11 No. 1, pp. 46–55.

Daly, S.R., Yilmaz, S., Christian, J.L., Seifert, C.M. and Gonzalez, R. (2012), “Design Heuristics in Engineering”, Vol. 101 No. 4, pp. 601–629.

Deliège, I. and Wiggins, G.A. (2006), *Musical Creativity: Multidisciplinary Research in Theory and Practice*, Psychology Press.

Dinar, M., Shah, J.J., Cagan, J., Leifer, L., Linsey, J.S., Smith, S.M. and Hernandez, N.V. (2015), “Empirical Studies of Designer Thinking: Past, Present, and Future”, *Journal of Mechanical Design*, Vol. 137 No. 2, p. 021101.

Doboli, A., Umbarkar, A., Subramanian, V. and Doboli, S. (2014), “Two experimental studies on creative concept combinations in modular design of electronic embedded systems”, *Design Studies*, Elsevier Ltd, Vol. 35 No. 1, pp. 80–109.

Dorst, K. and Cross, N. (2001), “Creativity in the design process: co-evolution of problem–solution”, *Design Studies*, Vol. 22 No. 5, pp. 425–437.

Dunn, O.J. (1964), “Multiple Comparisons Using Rank Sums”, *Technometrics*, Vol. 6 No. 3, pp. 241–252.

Eberle, B. (1996), *Scamper on: Games for Imagination Development*, Prufrock Press Inc.

Eisler, H. and Ekman, G. (1959), “A mechanism of subjective similarity”, *Nordisk Psykologi*, Taylor & Francis, Vol. 11 No. 1, pp. 1–10.

Estes, Z. (2003a), “Attributive and relational processes in nominal combination”, *Journal of Memory and Language*, Vol. 48 No. 2, pp. 304–319.

Estes, Z. (2003b), “A tale of two similarities: Comparison and integration in conceptual combination”, *Cognitive Science*, Vol. 27 No. 6, pp. 911–921.

Estes, Z. and Glucksberg, S. (2000), “Interactive property attribution in concept combination.”, *Memory & Cognition*, Vol. 28 No. 1, pp. 28–34.

Estes, Z., Golonka, S. and Jones, L.L. (2011), “Thematic Thinking”, *Psychology of Learning and Motivation - Advances in Research and Theory*, pp. 249–294.

Estes, Z. and Hasson, U. (2004), “The importance of being nonalignable: a critical test of the structural alignment theory of similarity.”, *Journal of Experimental Psychology. Learning, Memory, and Cognition*, Vol. 30 No. 5, pp. 1082–92.

Estes, Z. and Ward, T.B. (2002), “The emergence of novel attributes in concept modification”, *Creativity Research Journal*, Vol. 14 No. 2, pp. 253–266.

Fauconnier, G. and Turner, M. (1998), “Conceptual Integration Networks”, *Cognitive*

Science, Vol. 22 No. 2, pp. 133–187.

Fauconnier, G. and Turner, M. (2003), "Conceptual blending, form and meaning", *Recherches En Communication*, Vol. 19 No. 2003, pp. 57–86.

Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. (2007), "G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences", *Behavior Research Methods*, Springer, Vol. 39 No. 2, pp. 175–191.

Fidler, F., Singleton Thorn, F., Barnett, A., Kambouris, S. and Kruger, A. (2018), "The Epistemic Importance of Establishing the Absence of an Effect", *Advances in Methods and Practices in Psychological Science*, Vol. 1 No. 2, pp. 237–244.

Field, A. (2009), "Discopering Statistics Using SPSS, Thrid Edition", Sage Publications.

Finke, R.A. and Slayton, K. (1988), "Explorations of creative visual synthesis in mental imagery", *Memory & Cognition*, Springer-Verlag, Vol. 16 No. 3, pp. 252–257.

Finke, R.A., Ward, T.B. and Smith, S.M. (1992), "Creative cognition: Theory, research, and applications", *Computers & Mathematics with Applications*, Vol. 32, p. 128.

Fleetwood, S. (2014), "Bhaskar and Critical Realism", *The Oxford Handbook of Sociology, Social Theory, and Organization Studies*, Oxford University Press, p. 219.

Fleming, L. (2001), "Recombinant Uncertainty in Technological Search", Vol. 47 No. 1, pp. 117–132.

French, M.J. (1998), *Conceptual Design for Engineers*, Springer London, London, available at:https://doi.org/10.1007/978-1-4471-3627-9.

Frey, D.D., Herder, P.M., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K. and Clausing, D.P. (2009), "The Pugh Controlled Convergence method: Model-based evaluation and implications for design theory", *Research in Engineering Design*, Springer, Vol. 20 No. 1, pp. 41–58.

Fried, E.I. (2021), "Theories and Models: What They Are, What They Are for, and What They Are About", *Psychological Inquiry*, Routledge, Vol. 31 No. 4, pp. 336–344.

Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C. and Wood, K. (2013), "The Meaning of 'Near' and 'Far': The Impact of Structuring Design Databases and the Effect of Distance of Analogy on Design Output", *Journal of Mechanical Design*, Vol. 135 No. 2, p. 021007.

Gabora, L. (2004), "Creative Thought as a nonDarwinian Evolutionary Process", Vol. 39 No. 4, pp. 262–283.

Gagné, C.L. (2001), "Relation and lexical priming during the interpretation of noun–noun combinations.", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 27 No. 1, pp. 236–254.

Gagné, C.L. and Shoben, E.J. (1997), "Influence of Thematic Relations on the Comprehension of Modifier–Noun Combinations", *Journal of Experimental Psychology. Learning, Memory, and Cognition*, American Psychological Association, Vol. 23 No. 1, pp. 71–87.

Gentner, D. and Brem, S. (1999), “Is snow really like a shovel? Distinguishing similarity from thematic relatedness”, *Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society*, pp. 179–184.

Gentner, D. and Gunn, V. (2001), “Structural alignment facilitates the noticing of differences”, *Memory & Cognition*, Vol. 29 No. 4, pp. 565–577.

Gentner, D. and Markman, A.B. (1994), “Structural alignment in comparison: no difference without similarity”, *Psychological Science*, Vol. 5 No. 6, pp. 423–430.

Gentner, D. and Markman, A.B. (1995), “Similarity Is Like Analogy: Structural Alignment in Comparison”, *Similarity in Language, Thought and Perception*, pp. 111–147.

Gentner, D. and Markman, A.B. (1997), “Structure mapping in analogy and similarity”, *American Psychologist*, Vol. 52 No. 1, pp. 45–56.

Gentner, D. and Smith, L.A. (2012), *Analogical Reasoning*, *Encyclopedia of Human Behavior*, 2nd ed., Vol. 1, Elsevier Inc., available at:https://doi.org/10.1016/B978-0-12-375000-6.00022-7.

Gero, J.S. (1990), “Design Prototypes: A Knowledge Representation Schema for Design”, *AI Magazine*, Vol. 11 No. 4, p. 26.

Gero, J.S. (2000), “Computational Models of Innovative and Creative Design Processes”, *Technological Forecasting and Social Change*, Vol. 64 No. 2–3, pp. 183–196.

Gero, J.S. and Kannengiesser, U. (2004), “The situated function–behaviour–structure framework”, *Design Studies*, Vol. 25 No. 4, pp. 373–391.

Gero, J.S. and Kannengiesser, U. (2014), “The Function-Behaviour-Structure Ontology of Design”, *An Anthology of Theories and Models of Design*, Springer London, London, pp. 263–283.

Gerrig, R.J. and Bortfeld, H. (1999), “Sense Creation in and out of Discourse Contexts”, *Journal of Memory and Language*, Vol. 41 No. 4, pp. 457–468.

Ghasemi, A. and Zahediasl, S. (2012), “Normality tests for statistical analysis: A guide for non-statisticians”, *International Journal of Endocrinology and Metabolism*, Vol. 10 No. 2, pp. 486–489.

Gibbs, R.W. (2001), “Making Good Psychology Out of Blending Theory”, *Cognitive Linguistics*, Vol. 11 No. 3–4, pp. 347–358.

Gill, A.S., Tsoka, A.N. and Sen, C. (2019), “Dimensions of product similarity in design by analogy: An exploratory study”, *Proceedings of the ASME Design Engineering Technical Conference*, Vol. 7, American Society of Mechanical Engineers (ASME), available at:https://doi.org/10.1115/DETC2019-98252.

Glier, M.W., Tsenn, J., Linsey, J.S. and McAdams, D.A. (2014), “Evaluating the directed intuitive approach for bioinspired design”, *Journal of Mechanical Design, Transactions of the ASME*, Vol. 136 No. 7, pp. 1–9.

Goel, A.K. and Helms, M. (2014), “Theories, Models, Programs, and Tools of Design: Views from Artificial Intelligence, Cognitive Science, and Human-Centered Computing”, *An*

Anthology of Theories and Models of Design, pp. 173–195.

Goel, V. (1995), *Sketches of Thought*, The MIT Press, available at:https://doi.org/10.7551/mitpress/6270.001.0001.

Goldschmidt, G. and Smolkov, M. (2006), "Variances in the impact of visual stimuli on design problem solving performance", *Design Studies*, Vol. 27, pp. 549–569.

Goldstone, R.L. and Son, J.Y. (2012), *Similarity*, *The Oxford Handbook of Thinking and Reasoning*, available at:https://doi.org/10.1093/oxfordhb/9780199734689.013.0010.

Golonka, S. and Estes, Z. (2009), "Thematic relations affect similarity via commonalities.", *Journal of Experimental Psychology. Learning, Memory, and Cognition*, Vol. 35 No. 6, pp. 1454–1464.

Gonçalves, M. and Cash, P. (2021), "The life cycle of creative ideas: Towards a dual-process theory of ideation", *Design Studies*, Vol. 72, pp. 1–33.

Grant, M.J. and Booth, A. (2009), "A typology of reviews: An analysis of 14 review types and associated methodologies", *Health Information and Libraries Journal*, Vol. 26 No. 2, pp. 91–108.

Guba, E.G. and Lincoln, Y.S. (2014), "Competing paradigms in qualitative research", pp. 1–22.

Guilford, J.P. (1968), *Intelligence, Creativity, and Their Educational Implications*, Edits Pub.

Han, J., Shi, F., Chen, L. and Childs, P.R. (2018a), "A computational tool for creative idea generation based on analogical reasoning and ontology", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 32 No. 4, pp. 462–477.

Han, J., Shi, F., Chen, L. and Childs, P.R.N. (2018b), "The Combinator-a computer-based tool for creative idea generation based on a simulation approach", *Design Science*, Vol. 4, pp. 1–34.

Haslbeck, J., Ryan, O., Robinaugh, D., Waldorp, L. and Borsboom, D. (2019), "Modeling Psychopathology: From Data Models to Formal Theories", pp. 1–38.

Hay, L., Duffy, A.H.B., Grealy, M., Tahsiri, M., McTeague, C. and Vuletic, T. (2020), "A novel systematic approach for analysing exploratory design ideation", *Journal of Engineering Design*, Taylor & Francis, Vol. 31 No. 3, pp. 127–149.

Hay, L., Duffy, A.H.B., McTeague, C., Pidgeon, L.M., Vuletic, T. and Grealy, M. (2017a), "Towards a shared ontology: A generic classification of cognitive processes in conceptual design", *Design Science*, Vol. 3, p. e7.

Hay, L., Duffy, A.H.B., McTeague, C., Pidgeon, L.M., Vuletic, T. and Grealy, M. (2017b), "A systematic review of protocol studies on conceptual design cognition: Design as search and exploration", *Design Science*, Vol. 3 No. Visser 2004, p. e10.

Hebart, M.N., Zheng, C.Y., Pereira, F. and Baker, C.I. (2020), "Revealing the multidimensional mental representations of natural objects underlying human similarity judgements", *Nature Human Behaviour*, Nature Publishing Group, Vol. 4 No.

11, pp. 1173–1185.

Herrmann, D.J. and Chaffin, R. (1986), "Comprehension of semantic relations as a function of the definitions of relations", *Human Memory and Cognitive Capabilities: Mechanisms and Performance*, Elsevier and North-Holland Amsterdam, pp. 311–319.

Hill, M.R. (1984), "Epistemology, Axiology, and Ideology in Sociology", *Social Thought and Research*, available at:https://doi.org/10.17161/STR.1808.4967.

Hodgetts, C.J., Hahn, U. and Chater, N. (2009), "Transformation and alignment in similarity", *Cognition*, Elsevier, Vol. 113 No. 1, pp. 62–79.

Holtz, P. and Odağ, Ö. (2020), "Popper was not a Positivist: Why Critical Rationalism Could be an Epistemology for Qualitative as well as Quantitative Social Scientific Research", *Qualitative Research in Psychology*, Routledge, Vol. 17 No. 4, pp. 541–564.

Honke, G. and Kurtz, K.J. (2019), "Similarity is as similarity does? A critical inquiry into the effect of thematic association on similarity", *Cognition*, Elsevier, Vol. 186 No. June 2018, pp. 115–138.

Howard-Jones, P.A. (2002), "A dual-state model of creative cognition for supporting strategies that foster creativity in the classroom", *International Journal of Technology and Design Education*, Vol. 12 No. 3, pp. 215–226.

Howard, T.J., Culley, S.J. and Dekoninck, E. (2008), "Describing the creative design process by the integration of engineering design and cognitive psychology literature", *Design Studies*, Vol. 29 No. 2, pp. 160–180.

Jameson, J., Gentner, D., Day, S., Christie, S., Colhoun, J. and Bartels, D. (2005), "Clarifying the Role of Alignability in Similarity Comparisons", *Twenty-Seventh Annuam Meeting of the Cognitive Science Society*, pp. 1048–1053.

Jang, S. (2014), "The Effect of Image Stimulus on Conceptual Combination in the Design Idea Generation Process", *Archives of Design Research*, Vol. 27 No. 4, pp. 59–73.

Jang, S.H., Oh, B., Hong, S. and Kim, J. (2019), "The effect of ambiguous visual stimuli on creativity in design idea generation", *International Journal of Design Creativity and Innovation*, Taylor & Francis, Vol. 7 No. 1–2, pp. 70–98.

Jansson, D.G. and Smith, S.M. (1991), "Design fixation", *Design Studies*, Vol. 12 No. 1, pp. 3–11.

Jarecki, J.B., Tan, J.H. and Jenny, M.A. (2020), "A framework for building cognitive process models", *Psychonomic Bulletin and Review*, Psychonomic Bulletin & Review, Vol. 27 No. 6, pp. 1218–1229.

Jin, Y. and Benami, O. (2010), "Creative patterns and stimulation in conceptual design", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 24 No. 02, p. 191.

Jin, Y. and Chusilp, P. (2006), "Study of mental iteration in different design situations", *Design Studies*, Vol. 27 No. 1, pp. 25–55.

Jones, M. and Love, B.C. (2007), "Beyond common features: The role of roles in determining similarity", *Cognitive Psychology*, Vol. 55 No. 3, pp. 196–231.

Kan, J.W.T. and Gero, J.S. (2011), "Comparing Designing Across Different Domains : an Exploratory Case Study", *Design*, No. August.

Kim, J. (2011), *Modeling Cognitive and Affective Processes of Designers in the Early Stages of Design: Mental Categorization of Information Processing*.

Kim, J.E., Bouchard, C., Omhover, J.F. and Aoussat, A. (2010), "Towards a model of how designers mentally categorise design information", *CIRP Journal of Manufacturing Science and Technology*, CIRP, Vol. 3 No. 3, pp. 218–226.

Kokotovich, V. and Purcell, T. (2000), "Mental synthesis and creativity in design: an experimental examination", *Design Studies*, Vol. 21 No. 5, pp. 437–449.

Kozbelt, A., Beghetto, R.A. and Runco, M.A. (2010), "Theories of creativity", *The Cambridge Handbook of Creativity*, Vol. 2, pp. 20–47.

Krippendorff, K. (2018), *Content Analysis: An Introduction to Its Methodology*, Sage publications.

Kruger, C. (1999), *Cognitive Stategies in Industrial Design Engineering*, Delft University of Technology.

Kruger, C. and Cross, N. (2006), "Solution driven versus problem driven design: strategies and outcomes", *Design Studies*, Vol. 27 No. 5, pp. 527–548.

Krumhansl, C.L. (1978), "Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density.", *Psychological Review*, Vol. 85 No. 5, pp. 445–463.

Laerd Statistics. (2015a), "Statistical tutorials and software guides", *Statistical Tutorials and Software Guides*, available at: https://statistics.laerd.com/.

Laerd Statistics. (2015b), "Multiple regression using SPSS Statistics", *Statistical Tutorials and Software Guides*, available at: https://statistics.laerd.com/.

Laerd Statistics. (2015c), "Wilcoxon signed-rank test using SPSS Statistics", *Statistical Tutorials and Software Guide*, available at: https://statistics.laerd.com/.

Laerd Statistics. (2017), "Chi-square test of homogeneity (2 x C) using SPSS Statistics", *Statistical Tutorials and Software Guides2*, available at: https://statistics.laerd.com/.

Laerd Statistics. (2018), "Pearson's product-moment correlation using SPSS Statistics", *Statistical Tutorials and Software Guides*, available at: https://statistics.laerd.com/.

Lakoff, G. (2008), *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, University of Chicago press.

Landauer, T.K. (1997), "A Solution to Plato ' s Problem : The Latent Semantic Analysis Theory of Acquisition , Induction , and Representation of Knowledge A Solution to Plato ' s Problem : The Latent Semantic Analysis Theory of Acquisition , Induction , and Representation of Kno", Vol. 1 No. 2, pp. 1–72.

Liikkanen, L.A. (2010), *Design Cognition for Conceptual Design*, Aalto University.

Liikkanen, L.A. and Perttula, M. (2010), “Inspiring design idea generation: Insights from a memory-search perspective”, *Journal of Engineering Design*, Vol. 21 No. 5, pp. 545–560.

Liu, Y.C., Bligh, T. and Chakrabarti, A. (2003), “Towards an ‘ideal’ approach for concept generation”, *Design Studies*, Vol. 24 No. 4, pp. 341–355.

Lloyd, P., Lawson, B. and Scott, P. (1995), “Can concurrent verbalization reveal design cognition?”, *Design Studies*, Vol. 16 No. 2, pp. 237–259.

Lloyd, P. and Scott, P. (1994), “Discovering the design problem”, *Design Studies*, Vol. 15 No. 2, pp. 125–140.

Lubart, T. (2001), “Models of the creative process: Past, present and future”, *Creativity Research Journal*, Vol. 13 No. 3–4, pp. 295–308.

Luo, J., Sarica, S. and Wood, K.L. (2021), “Guiding data-driven design ideation by knowledge distance”, *Knowledge-Based Systems*, Elsevier B.V., Vol. 218, p. 106873.

Lynott, D. and Connell, L. (2010), “Embodied conceptual combination”, *Frontiers in Psychology*, Vol. 1 No. NOV, pp. 1–14.

Maguire, P., Maguire, R. and Cater, A.W.S. (2010), “The Influence of Interactional Semantic Patterns on the Interpretation of Noun-Noun Compounds”, *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, Vol. 36 No. 2, pp. 288–297.

Maher, M.L. (2000), “A model of co-evolutionary design”, *Engineering with Computers*, Vol. 16 No. 3–4, pp. 195–208.

Maher, M.L., Poon, J. and Boulanger, S. (1996), “Formalising Design Exploration as Co-evolution: A Combined Gene Approach”, *Advances in Formal Design Methods for CAD: Proceedings of the IFIP WG5.2 Workshop on Formal Design Methods for Computer-Aided Design*, pp. 3–30.

Markman, A.B. (1996), “Structural Alignment in Similarity and Difference Judgments”, *Psychonomic Bulletin & Review*, Vol. 3 No. 2, pp. 227–230.

Markman, A.B. (2013), *Knowledge Representation*, Psychology Press.

Markman, A.B. and Gentner, D. (1993a), “Structural alignment during similarity comparisons”, *Cognitive Psychology*.

Markman, A.B. and Gentner, D. (1993b), “Splitting the differences: A structural alignment view of similarity”, *Journal of Memory and Language*, Vol. 32 No. 4, pp. 517–535.

Markman, A.B. and Gentner, D. (1996), “Commonalities and differences in similarity comparisons”, *Memory & Cognition*, Vol. 24 No. 2, pp. 235–249.

Markman, A.B. and Rein, J.R. (2013), “The nature of mental concepts”, *The Oxford Handbook of Cognitive Psychology*, No. May 2018, pp. 321–345.

Markman, A.B. and Wisniewski, E.J. (1997), “Similar and different: The differentiation of basic-level categories.", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 23 No. 1, pp. 54–70.

Marr, D. (1982), *Vision*, The MIT Press, available at:https://doi.org/10.7551/mitpress/9780262514620.001.0001.

Marr, D. and Poggio, T. (1976), *From Understanding Computation to Understanding Neural Circuitry*.

McNeill, T., Gero, J.S., Warren, J. and Neill, T.M. (1998), "Understanding conceptual electronic design using protocol analysis", *Research in Engineering Design*, Vol. 10 No. 3, pp. 129–140.

McTeague, C.P., Duffy, A., Hay, L., Vuletic, T., Campbell, G., Choo, P.L. and Grealy, M. (2018), "Insights into design concept similarity judgements", *Proceedings of International Design Conference, DESIGN*, Vol. 5, pp. 2087–2098.

Medin, D.L., Goldstone, R.L. and Gentner, D. (1990), "Similarity involving attributes and relations: Judgments of similarity and difference are not inverses", *Psychological Science*, Vol. 1 No. 1, pp. 64–69.

Merriam-Webster. (n.d.). "Ontology", *Merriam-Webster.Com Dictionary*, available at: https://www.merriam-webster.com/dictionary (accessed 7 June 2021a).

Merriam-Webster. (n.d.). "Epistemology", *Merriam-Webster.Com Dictionary*, available at: https://www.merriam-webster.com/dictionary/epistemology (accessed 7 June 2021b).

Merriam-Webster. (n.d.). "Axiology", *Merriam-Webster.Com Dictionary*, available at: https://www.merriam-webster.com/dictionary/axiology (accessed 7 June 2021c).

Miller, G.A. (1995), "WordNet: a lexical database for English", *Communications of the ACM*, Vol. 38, ACM New York, NY, USA, pp. 39–41.

Mirman, D. and Graziano, K.M. (2012), "Individual differences in the strength of taxonomic versus thematic relations", Vol. 141 No. 4, pp. 601–609.

Mirman, D., Landrigan, J.-F. and Britt, A.E. (2017), "Taxonomic and Thematic Semantic Systems Taxonomic and Thematic Semantic Systems", *Psychological Bulletin*, Vol. 143 No. 5, pp. 499–520.

Mobley, M.I., Doares, L.M. and Mumford, M.D. (1992), "Process Analytic Models of Creative Capacities: Evidence for the Combination and Reorganization Process", *Creativity Research Journal*, Vol. 0419 No. 2, pp. 125–155.

Mumford, M.D., Baughman, W.A., Maher, M.A., Costanza, D.P. and Supinski, E.P. (1997), "Process-based measures of creative problem-solving skills: IV. Category combination", *Creativity Research Journal*, Vol. 10 No. 1, pp. 33–44.

Mumford, M.D., Medeiros, K.E. and Partlow, P.J. (2012), "Creative thinking: Processes, strategies, and knowledge", *Journal of Creative Behavior*, Vol. 46 No. 1, pp. 30–47.

Mumford, M.D., Mobley, M.I., Reiter-Palmon, R., Uhlman, C.E., Doares, L.M., Reiter Palmon, R., Uhlman, C.E., *et al.* (1991), "Process Analytic Models of Creative Capacities", *Creativity Research Journal*, Vol. 4 No. 2, pp. 91–122.

Murphy, G. (2004), *The Big Book of Concepts*, MIT press.

Murphy, G.L. (2001), “Causes of taxonomic sorting by adults: a test of the thematic-to-taxonomic shift.”, *Psychonomic Bulletin & Review*, Vol. 8 No. 4, pp. 834–839.

Murphy, G.L. (2010), “What are categories and concepts?”, *The Making of Human Concepts*, Vol. 6, Oxford University Press, pp. 11–28.

Nagai, Y. and Taura, T. (2006), “Formal Description of Concept-synthesizing Process for Creative Design”, *Design Computing and Cognition’06*, pp. 443–460.

Nagai, Y., Taura, T. and Mukai, F. (2009), “Concept blending and dissimilarity: factors for creative concept generation process”, *Design Studies*, Elsevier Ltd, Vol. 30 No. 6, pp. 648–675.

Nandy, A., Dong, A. and Goucher-Lambert, K. (2020), “A comparison of vector and network-based measures for assessing design similarity”, *Proceedings of the ASME Design Engineering Technical Conference*, Vol. 8, available at:https://doi.org/10.1115/detc2020-22424.

Nandy, A. and Goucher-Lambert, K. (2021), “Aligning human and computational evaluations of functional design similarity”, *Proceedings of the ASME Design Engineering Technical Conference*, Vol. 6, available at:https://doi.org/10.1115/DETC2021-71905.

Nandy, A. and Goucher-Lambert, K. (2022), “Do Human and Computational Evaluations of Similarity Align? An Empirical Study of Product Function”, *Journal of Mechanical Design*, Vol. 144 No. 4, available at:https://doi.org/10.1115/1.4053858.

Nelson, R.R. and Winter, S.G. (1977), “In search of useful theory of innovation”, *Research Policy*, Vol. 6 No. 1, pp. 36–76.

Nguyen, T.A. and Zeng, Y. (2014), “A physiological study of relationship between designer’s mental effort and mental stress during conceptual design”, *Computer-Aided Design*, Vol. 54, pp. 3–18.

Nijstad, B.A. and Sroebe, W. (2006), “How the Group Affects the Mind:A Cognitive Model of Idea Generation in Groups”, *Personality and Social Psychology Review*, Vol. 10 No. 3, pp. 186–213.

Osborn, A.F. (1957), *Applied Imagination; Principles and Procedures of Creative Thinking, by Alex. F. Osborn*.

Öztuna, D., Elhan, A.H. and Tüccar, E. (2006), “Investigation of four different normality tests in terms of type 1 error rate and power under different distributions”, *Turkish Journal of Medical Sciences*, The Scientific and Technological Research Council of Turkey, Vol. 36 No. 3, pp. 171–176.

Palmer, S. (1978), “Fundamental aspects of cognitive representation”.

Paton, B. and Dorst, K. (2011), “Briefing and reframing: A situated practice”, *Design Studies*, Elsevier Ltd, Vol. 32 No. 6, pp. 573–587.

Pearson, D.G., Logie, R.H. and Gilhooly, K.J. (1999), "Verbal Representations and Spatial Manipulation During Mental Synthesis", *European Journal of Cognitive Psychology*, Vol. 11 No. 3, pp. 295–314.

Pedgley, O.F. (1999), *Industrial Designers' Attention to Materials and Manufacturing Processes: Analyses at Macroscopic and Microscopic Levels*, *Department of Design and Technology*, available at: http://www.theeuropeanlibrary.org/tel4/record/2000024073298.

Perttula, M. and Liikkanen, L.A. (2006), "Structural tendencies and exposure effects in idea generation", pp. 1–11.

Petre, M. (2004), "How expert engineering teams use disciplines of innovation", *Design Studies*, Vol. 25 No. 5, pp. 477–493.

Pilgrim, D. (2020), *Critical Realism for Psychologists*, *Critical Realism for Psychologists*, Routledge, available at:https://doi.org/10.4324/9780429274497.

Poldrack, R.A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Parker, D.S., *et al.* (2011), "The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience", *Frontiers in Neuroinformatics*, Vol. 5 No. September, pp. 1–11.

Ponterotto, J.G. (2005), "Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science", *Journal of Counseling Psychology*.

Popper, K. (1959), *The Logic of Scientific Discovery*, Routledge.

Popper, K. (1963), *Conjectures and Refutations: The Growth of Scientific Knowledge*, routledge.

Pugh, S. (1991), *Total Design - Integrated Methods for Successful Product Engineering*, Addison-Wesley Pub. Co.

Purcell, A.T. and Gero, J.S. (1996), "Design and other types of fixation", *Design Studies*, Vol. 17 No. 4 SPEC. ISS., pp. 363–383.

Qian, L. and Gero, J.S. (1992), "A Design Support System Using Analogy Based Reasoning", *Artificial Intelligence in Design*, pp. 795–813.

Qian, L. and Gero, J.S. (1996), "Function–behavior–structure paths and their role in analogy-based design", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 10 No. 4, pp. 289–312.

Ran, B. and Duimering, R.P. (2009), "Conceptual combination: models, theories, and controversies".

Ranawat, A. and Hölttä-Otto, K. (2009), "Four dimensions of design similarity", *Proceedings of the ASME Design Engineering Technical Conference*, Vol. 8, pp. 1069–1077.

Reich, Y. (1994), "Layered models of research methodologies", No. May 2021, pp. 263–274.

Reiter-palmon, R., Herman, A.E., Yammarino, F.J., Herman, A.E., Sacramento, C.A., Dawson, J.F., West, M.A., *et al.* (2015), "Creativity and cognitive processes: Multi-level linkages between individual and team cognition".

Robson, C. (2011), *Real World Research : A Resource for Users of Social Research Methods in Applied Settings*, Wiley.

Roozenburg, N.F.M. and Eekels, J. (1995), *Product Design : Fundamentals and Methods*, Wiley.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M. and Boyes-Braem, P. (1976), “Basic objects in natural categories”, *Cognitive Psychology*, Vol. 8 No. 3, pp. 382–439.

Rosenman, M.. and Gero, J.S. (1998), “Purpose and function in design: from the socio-cultural to the techno-physical”, *Design Studies*, Vol. 19 No. 2, pp. 161–186.

Rowe, P. (1991), *Design Thinking*, The MIT Press.

Runco, M.A. and Chand, I. (1995), “Cognition and creativity”, *Educational Psychology Review*, Vol. 7 No. 3, pp. 243–267.

Ryan, G.S. (2019), “Postpositivist critical realism: philosophy, methodology and method for nursing research”, *Nurse Researcher*, Vol. 27 No. 3, pp. 20–26.

Saunders, M.N.K., Lewis, P. and Thornhill, A. (2019), *Research Methods for Business Students Ebook*, Pearson Education.

Scheel, A.M., Tiokhin, L., Isager, P.M. and Lakens, D. (2021), “Why Hypothesis Testers Should Spend Less Time Testing Hypotheses”, *Perspectives on Psychological Science*, Vol. 16 No. 4, pp. 744–755.

Schoenmakers, W. and Duysters, G. (2010), “The technological origins of radical inventions”, *Research Policy*, Elsevier B.V., Vol. 39 No. 8, pp. 1051–1059.

Scott, G.M., Lonergan, D.C. and Mumford, M.D. (2005), “Conceptual Combination: Alternative Knowledge Structures, Alternative Heuristics”, *Creativity Research Journal*, Vol. 17 No. 1, pp. 79–98.

Seitamaa-hakkarainen, P., Huotilainen, M., Groth, C. and Hakkarainen, K. (2014), “The promise of cognitive neuroscience in design studies”, *Proceedings of DRS*, pp. 834–846.

Seitamaa-hakkarainen, P., Huotilainen, M. and Mäkelä, M. (2016), “How can neuroscience help understand design and craft activity ? The promise of cognitive neuroscience in design studies”, Vol. 9 No. 1, pp. 1–16.

Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, *Experimental and Quasi-Experimental Design for Causual Inference*, Boston: Houghton Mifflin,.

Shiffrin, R.M. and Atkinson, R.C. (1971), “The control of short-term memory”, *Scientific American*, Vol. 225, pp. 82–90.

Sim, S.K. and Duffy, A.H.B. (2003), “Towards an ontology of generic engineering design activities”, *Research in Engineering Design*, Vol. 14 No. 4, pp. 200–223.

Simmons, S. and Estes, Z. (2008), “Individual differences in the perception of similarity and difference”, *Cognition*, Vol. 108 No. 3, pp. 781–795.

Simonton, D.K. (1999), *Origins of Genius: Darwinian Perspectives on Creativity*, Oxford University Press.

Sjoberg, L. (1972), "A cognitive theory of similarity", *Goteborg Psychological Reports*, Vol. 2 No. 10.

Smiley, S.S. and Brown, A.L. (1979), "Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age", *Journal of Experimental Child Psychology*, Elsevier, Vol. 28 No. 2, pp. 249–257.

Snodgrass, J.G. and Vanderwart, M. (1980), "A Standardized Set of 260 Pictures : Norms for Name Agreement , Image Agreement , Familiarity , and Visual Complexity", Vol. 6 No. 2, pp. 174–215.

Sosa, R. (2018), "Metrics to select design tasks in experimental creativity research", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, Vol. 0 No. 0, pp. 1–11.

Sowden, P.T., Pringle, A. and Gabora, L. (2015), "The shifting sands of creative thinking: Connections to dual-process theory", *Thinking and Reasoning*, Vol. 21 No. 1, pp. 40–60.

Spielman, R.M., Jenkins, W.J. and Lovett, M.D. (2020), "Psychology 2e", *Psychology 2e*, OpenStax, available at: https://openstax.org/books/psychology-2e/pages/1-introduction.

Stauffer, L.A. and Ullman, D.G. (1991), "Fundamental Processes of Mechanical Designers Based on Empirical Data", *Journal of Engineering Design*, Vol. 2 No. 2, pp. 113–125.

Štorga, M., Andreasen, M., Marjanović, D., Marioˇmarioštorga, M. and Marjanovi´c, D.M. (2008), "The design ontology: foundation for the design knowledge exchange and management", *Http://Dx.Doi.Org/10.1080/09544820802322557*, Taylor & Francis, Vol. 21 No. 4, pp. 427–454.

Storms, G. and Wisniewski, E.J. (2005), "Does the order of head noun and modifier explain response times in conceptual combination?", *Memory and Cognition*, Vol. 33 No. 5, pp. 852–861.

Suwa, M. and Tversky, B. (1997), "What do architects and students perceive in their design sketches? A protocol analysis", *Design Studies*, Vol. 18 No. 4, pp. 385–403.

Taura, T. (2016a), *Creative Design Engineering: Introduction to an Interdisciplinary Approach*, Academic Press, 2016.

Taura, T. (2016b), "Using Thematic Relations : Theory and Methodology of Concept Generation (Third Method)", *Creative Design Engineering*, Elsevier Inc., pp. 71–78.

Taura, T. (2016c), "The Blending Method : Theory and Methodology of Concept Generation (Second Method)", *Creative Design Engineering*, Elsevier Inc., pp. 57–69.

Taura, T. and Nagai, Y. (2013a), "A systematized theory of creative concept generation in design: First-order and high-order concept generation", *Research in Engineering Design*, Vol. 24 No. 2, pp. 185–199.

Taura, T. and Nagai, Y. (2013b), *Concept Generation for Design Creativity: A Systematized Theory and Methodology, Concept Generation for Design Creativity*, Vol. 9781447140, Springer London, London, available at:https://doi.org/10.1007/978-1-4471-4081-8.

Taura, T., Nagai, Y., Morita, J. and Takeuchi, T. (2007), "a Study on Design Creative Process Focused on Concept Combination Types in Comparison With Linguistic Interpretation", *Integration The Vlsi Journal*, No. August, pp. 1–12.

Tversky, A. (1977), "Features of Similarity", *American Psychologist*, Vol. 69 No. 5, pp. 379–399.

Tversky, A. and Gati, I. (1978), "Studies of similarity", *Cognition and Categorization*, Vol. 1 No. 1978, pp. 79–98.

Ullman, D.G. (2010), *The Mechanical Design Process*, *Mechanics of Materials*, McGraw-Hill Higher Education.

Ullman, D.G., Dietterich, T.G. and Stauffer, L. a. (1988), "A model of the mechanical design process based on empirical data", *Artificial Intelligence for Engineering, Design, Analysis and Manufacturing*, Vol. 2 No. 01, p. 33.

Umeda, Y., Takeda, H., Tomiyama, T. and Yoshikawa, H. (1990), "Function, behaviour, and structure".

Vasconcelos, L.A. and Crilly, N. (2016), "Inspiration and fixation: Questions, methods, findings, and challenges", *Design Studies*, Elsevier Ltd, Vol. 42, pp. 1–32.

Verstijnen, I., van Leeuwen, C., Goldschmidt, G., Hamel, R. and Hennessey, J. (1998a), "Sketching and creative discovery", *Design Studies*, Vol. 19 No. 4, pp. 519–546.

Verstijnen, I., van Leeuwen, C., Goldschmidt, G., Hamel, R. and Hennessey, J. (1998b), "Creative discovery in imagery and perception: combining is relatively easy, restructuring takes a sketch.", *Acta Psychologica*, Vol. 99 No. 2, pp. 177–200.

Verstijnen, I., van Leeuwen, C., Hamel, R. and Hennessey, J. (2000), "What Imagery Can't Do and Why Sketching Might Help", *Empirical Studies of the Arts*, Vol. 18 No. 2, pp. 167–182.

Visser, W. (2006), *The Cognitive Artifacts of Designing*, CRC Press.

Vuletic, T., Duffy, A., Hay, L., McTeague, C., Pidgeon, L. and Grealy, M. (2018), "The challenges in computer supported conceptual engineering design", *Computers in Industry*, Elsevier, 1 February.

Wahyuni, D. (2012), "The Research Design Maze: Understanding Paradigms, Cases, Methods and Methodologies", *Journal of Applied Management Accounting Research*, Vol. 10 No. 1, pp. 69–80.

Ward, T.B. and Kolomyts, Y. (2010), "Cognition and creativity", *The Cambridge Handbook of Creativity*, Cambridge University Press New York, NY, pp. 93–112.

Ward, T.B., Smith, S.M. and Vaid, J. (1997), *Creative Thought : An Investigation of Conceptual Structures and Processes*, American Psychological Association.

Wilkenfeld, M.J. and Ward, T.B. (2001), "Similarity and emergence in conceptual combination", *Journal of Memory and Language*, Vol. 45 No. 1, pp. 21–38.

Wilson, J.O., Rosen, D., Nelson, B.A. and Yen, J. (2010), "The effects of biological examples in idea generation", *Design Studies*, Elsevier Ltd, Vol. 31 No. 2, pp. 169–186.

Wisniewski, E.J. (1996), "Construal and Similarity in Conceptual Combination", Vol. 453, pp. 434–453.

Wisniewski, E.J. (1997a), "When concepts combine", *Psychonomic Bulletin & Review*, Vol. 4 No. 2, pp. 167–83.

Wisniewski, E.J. (1997b), "Conceptual combination: Possibilities and esthetics.", *Creative Thought: An Investigation of Conceptual Structures and Processes.*, American Psychological Association, Washington, pp. 51–81.

Wisniewski, E.J. (1998), "Property instantiation in conceptual combination.", *Memory & Cognition*, Vol. 26 No. 6, pp. 1330–47.

Wisniewski, E.J. (2000), "Similarity, alignment, and conceptual combination: Comment on Estes and Glucksberg", *Society*, Vol. 77 No. 1, pp. 189–194.

Wisniewski, E.J. (2001), "On the Necessity of Alignment: A reply to Costello and Keane (2001)", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 27 No. 1, pp. 272–277.

Wisniewski, E.J. and Bassok, M. (1999), "What makes a man similar to a tie? Stimulus compatibility with comparison and integration.", *Cognitive Psychology*, Vol. 39, pp. 208–238.

Wisniewski, E.J. and Love, B. (1998), "Relations versus properties in conceptual combination", *Journal of Memory and Language 38, 177–202 (1998)*, Vol. 38 No. 38, pp. 177–202.

Wisniewski, E.J. and Markman, A.B. (1993), "The Role of Structural Alignment in Conceptual Combination", *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*.

Wisniewski, E.J. and Middleton, E.L. (2002), "Of Bucket Bowls and Coffee Cup Bowls: Spatial Alignment in Conceptual Combination", *Journal of Memory and Language*, Vol. 46 No. 1, pp. 1–23.

Wood, W.H. and Agogino, A.M. (1996), "Case-based conceptual design information server for concurrent engineering", *CAD Computer Aided Design*, Elsevier Ltd, Vol. 28 No. 5, pp. 361–369.

Xu, X. and Paulson, L. (2013), "Concept Abstractness and the Representation of Noun-Noun Combinations", *Journal of Psycholinguistic Research*, Vol. 42 No. 5, pp. 413–431.

Yokoi, T. (1995), "The EDR electronic dictionary", *Communications of the ACM*, ACM New York, NY, USA, Vol. 38 No. 11, pp. 42–44.

Zec, S., Soriani, N., Comoretto, R. and Baldi, I. (2017), "High Agreement and High Prevalence: The Paradox of Cohen's Kappa.", *The Open Nursing Journal*, Vol. 11, pp.

211–218.

Zwicky, F. (1967), “The morphological approach to discovery, invention, research and construction”, *New Methods of Thought and Procedure*, Springer, pp. 273–297.

APPENDICES

APPENDIX 1 – INTRODUCTORY LITERATURE

APPENDIX 1A – RESEARCH METHODS FOR STUDYING DESIGN COGNITION

There are a range of techniques available to study design cognition generally. Previous reviews have addressed general empirical methods for studying design cognition (Coley *et al.*, 2007; Cross, 1990; Dinar *et al.*, 2015), as well as methods specifically for longitudinal studies (Pedgley, 1999) physiological measures (Kim, 2011) and neurological measures (Seitamaa-hakkarainen *et al.*, 2014, 2016).

Specific research methods are briefly assessed in terms of their usefulness for studying design cognition generally. Table A1- 1 describes research methods associated with ethnographic, behavioural and physiological studies, and also provides the pros and cons of each method. Neurological methods are not reviewed, but see (Seitamaa-hakkarainen *et al.*, 2016) for a review.

Table A1- 1 – Research methods for studying design cognition

Study type	Example	Description
Observations of designers at work	(Visser, 2006)	Qualitative research method which provides a description and interpretation of the social and cultural factors influencing a social group (Robson, 2011). Conducted in the participant's natural setting, such as at work. Can be paired with interviews.
Interviews with designers	(Petre, 2004) (Paton and Dorst, 2011)	Spoken exchanges between interviewer and interviewee. Can be structure or unstructured.
Case studies	(Rowe, 1991)	An empirical study which investigates a specific phenomenon within its natural context. Difficult to assess the generality of conclusions or their causal inferences (Dinar *et al.*, 2015)
Diary method	(Ball 1990) from Pedgeley (1999)	Participant is instructed to record their thoughts in a diary so that the experimenter might gain insight into their thoughts. Participants can act naturally, without direct influence from the experimenter.
Protocol study	(Suwa and Tversky, 1997) (McNeill *et al.*, 1998) (Chan, 1990)	Requires participants to either 'think-aloud' during a task or to give a retrospective account of their behaviour during the task. Allows the researcher to infer the thought processes of the participant
Predictors of cognitive ability	(Verstijnen et. al, 1998)	Correlations between an individual's cognitive ability and their task performance can identify whether certain cognitive processes are important in successful completion of that task. E.g. memory may be involved in a task, but memory capacity may not be a relevant determinant of performance in a task.
Controlled (Laboratory) Experiments	(Christensen and Schunn, 2007) (Purcell and Gero, 1996) (Goldschmidt and Smolkov, 2006)	Controlled experiments can use a number of various data gathering techniques in order to test the effect of altering variables and experimental conditions.
ECG (Cardiovascular system)	(Nguyen and Zeng, 2014)	Measures heart rate, including interval duration and heart rate variability. Can be used to infer certain aspects of an individual's emotional and cognitive state.
Eye-tracking	(Kim, 2011)	Can determine gaze position, visual fixations, search patterns. Provides information on the perceptual activities of the individual.

NB: Additional physiological measurements have been identified but are not included in the table e.g. Electrodermal activity (EDA), Body temperature, Pupillometric response, Electrooculography (EOG), Electromyography (EMG)

Appendix 2 – Combination literature

Appendix 2A – Combination constructs in design

Table A2-1 - List of protocol codes for combination constructs

Source	**Phenomena**	**Description**	**Goal / control**	**Input**	**Output**	**Subordinates**
(Kruger, 1999; Kruger and Cross, 2006)	Assemble	Partial solutions are combined	Goal: "assembling of partial solutions and solving conflicts between them"	Partial solutions (previously generated, described by: attributes, values, constraints)	Final design (iteration possible)	Consistency check Combination Repair
(Sim and Duffy, 2003)	Composing	"The combination of ideas/concepts through association of ideas/concepts that satisfy overall function"	Goal: "Combine ideas/concepts through association of ideas/ concepts that satisfy overall function"	Domain knowledge Combination tables, function modules, ideas / concepts	Concepts or modules that satisfy the overall functions	N/A
(Gero and Kannengiesser, 2004, 2014)	Focusing and interpretation processes	Interpretation of external F, B or S Focusing (of F, B or S)	Goals are derived from functions. The expected world contains goals for desired changes to the external world.	F^e, B^e or S^e F^i B^i or S^i	F^i B^i or S^i F^i B^i or S^i	N/A N/A
(Jin and Chusilp, 2006)	Compose	"Compose involves the evolution of initial design ideas into identifiable design concepts" "The designer combines entity e1 and e2, and then transforms them into an evolved entity e3".	Design goals are formulated by the 'analyze' activity. The 'compose' activity is controlled by problem requirements and constraints.	"Entities" (Previous concepts or newly generated ideas)	Evolved entity	Associate Transform
(Daly *et al.*, 2012)	Synthesise	"Took two or more previous concepts and merged them"	N/A	Concepts (undefined)		N/A

Table A2-2 - List of outcome-based experimental studies of design combination published in journal articles

	Goal		Input					Output	
Source	**Goal variables**	**T**	**# stim / trial**	**Type of input**	**Stimuli bank**	**Manipulation**	**Control**	**Output**	**DV**
(Nagai *et al.*, 2009)	'Design a new concept'	10	Pairs of concepts 2 trials	Two object-pairs (Noun-noun word pairs), e.g. "ship-box" "piano-guitar"	2 pairs Ship-guitar Desk-elevator	N/A	- associative effectiveness - taxonomic category membership - number of commonalities and differences	A new concept (sketch and explanation sentence)	Originality Practicality Emergent features Combination type
(Doboli *et al.*, 2014)	Develop and describe a novel electronic embedded system that is useful, use as many devices from the list as possible	10	12+ stimuli (varies per group) 1 trial	'Building blocks for modular design solutions', e.g., 'gps', 'motor', 'memory' (words)	All presented at once	- Salience, - relatedness, - number	N/A	'Novel electronic embedded system'	- Frequency of combination types - Novelty, - quality - usefulness
(Jang, 2014)	Unclear	5	Pairs of concepts 40 trials	7 concepts from 13 categories, including natural and artificial objects (Represented as words or pictures)	40 randomly presented pairs	- Similarity - Taxonomic category (same/dif) - Representation modality	N/A	Instructions unclear, but can be called 'idea sketch'	Creativity
(Jang *et al.*, 2019)	Task: 'to design idea sketches' Encouraged to sketch as many concepts as possible.	5	Pairs of concepts 24 trials	Objects (Sketches of varying 'abstraction') Stimuli from a bank of 260 pictures.	24 noun pairs, 12 from the same category, 12 from different.	- Similarity (same or different category) - Abstraction	Visual complexity	"To design idea sketches" 2 idea sketches per pair	-Abstractness -Elaboration -Fluency -Resistance to premature closure -Originality

T = time given to produce one output, # stim = number of stimuli presented, DV = dependent variable

Table A2-3 - List of cognitive models of creative output production in conceptual design

Source(s)	**Model / study description**	**Combination process**	**Subordinate constructs**	**Empirical evidence**	**Mod**	**Pro**	**Out**
(Kruger, 1999; Kruger and Cross, 2006)	An expertise model of product design	Combination is the process 'assembly', distinct from and complementary to generation.	Consistency check Combination Repair	Model developed using the CommonKADS conceptual modelling language and evaluated via protocol analysis.	Y	Y	-
(Sim and Duffy, 2003)	A taxonomy of design activities.	Combination occurs via the 'composing' activity. One of three activities (alongside generating and associating) that form a compound activity 'concept generation'.	N/A	Design activities derived from a literature review and evaluated via protocol analysis.	Y	Y	-
(Gero and Kannengiesser, 2004)	A situated framework of designing	Combination is part of a class of processes that trigger reformulation.	N/A	The FBS ontology and its associated frameworks (Gero, 1990; Gero and Kannengiesser, 2004) have received validation through numerous protocol studies. See Gero and Kannengiesser (2014) for examples.	Y	Y	-
(Jin and Chusilp, 2006)	A cognitive activity model of conceptual design	In the initial conceptualisation of the model, the 'compose' activity was purposefully included by the authors to facilitate the study of "how iteration interacts with idea generation and evolution" (p.31).	Associate Transform	Model proposed based on past work (Jin and Benami, 2010) and verified and adjusted with data from protocol studies of 16 systems engineering or mechanical engineering students.	Y	Y	-
(Taura, 2016a; Taura and Nagai, 2013b, 2013a)	A systematised theory of concept generation	Three combination processes are proposed as methods for concept generation	Each combination process has a series of steps	The 'theory' has been developed using protocol analysis, experiments, logical reasoning and computational modelling. The three methods have been validated in experiments (Nagai *et al.*, 2009)	Y	Y	
(Stauffer and Ullman, 1991; Ullman *et al.*, 1988)	Mechanical design process	No explicit combination component. Could be accounted for through iterative generate-and-test or generate-and-improve cycles.	N/A	No explicit combination component. Could be accounted for through iterative generate-and-test or generate-and-improve cycles.	Y	Y	-

Source(s)	Model / study description	Combination process	Subordinate constructs	Empirical evidence	Mod	Pro	Out
(Chan, 1990)	Architectural design problem solving	No explicit combination component. Could be accounted for through iterative generate-and-test cycles, where prior solutions are retained for use in subsequent cycles.	N/A	No explicit combination component. Could be accounted for through iterative generate-and-test cycles, where prior solutions are retained for use in subsequent cycles.	Y	Y	-
(Benami and Jin, 2002; Jin and Benami, 2010)	Conceptual design	No explicit combination component. A mental synthesis process (Finke *et al.*, 1992) hypothesised to exist but not identifiable in protocol.	N/A	No explicit combination component. A mental synthesis process (Finke *et al.*, 1992) hypothesised to exist but not identifiable in protocol.	Y	Y	-
(Kim *et al.*, 2010)	How designers categorise information	No explicit combination component.	N/A	No explicit combination component.	Y	Y	-
(Liikkanen and Perttula, 2010)	Model -L Idea generation (specifically memory search)	The model explicitly does forgoes a 'recomposition' mechanism. Combination is possible by using prior ideas as inputs to the idea generation process.	Micro-level of Model-L		Y	Y	Y
(Daly *et al.*, 2012)	A list of design heuristics in engineering concept generation	'Synthesize' is one of four 'general approaches' used to generate new concepts from existing ideas. More specific heuristics such as 'merge functions with same energy source' or 'unify' were also identified.	N/A	The list of heuristics was created by coding protocol data in a mixed deductive (using a pre-defined exemplar list) and inductive (extracting new phenomena from the data) approach	-	Y	-
(Chan and Schunn, 2015)	Study of mental iteration in conceptual combination	Instances of combination identified through retrospective, genealogical analysis of OpenIDEO database.	N/A	Combination defined by concepts and their antecedents	-	-	Y
(Nagai *et al.*, 2009)	Factors for creative concept generation processes	Combination elicited by experiment design	N/A	Combination defined by concepts (task outputs) and their antecedents (stimuli)	-	-	Y
(Doboli *et al.*, 2014)	Concept combination in electronic systems	Combination elicited by experiment design	N/A	Combination defined by concepts (task outputs) and their antecedents (stimuli)	-	-	Y

Source(s)	Model / study description	Combination process	Subordinate constructs	Empirical evidence	Mod	Pro	Out
(Jang, 2014)	The effect of image stimulus on conceptual combination in design idea generation	Combination elicited by experiment design	N/A	Combination defined by concepts (task outputs) and their antecedents (stimuli)	-	-	Y
(Jang *et al.*, 2019)	The effect of image stimulus on conceptual combination in design idea generation	Combination elicited by experiment design	N/A	Combination defined by concepts (task outputs) and their antecedents (stimuli)	-	-	Y

APPENDIX 3 – MATERIALS AND METHODS

APPENDIX 3A – PARTICIPANT DEMOGRAPHICS

Table A3-1 – Demographic data for Sim-P1

ID	Age	UG / Masters	PG	Year of study	Institution
P1	24	MEng Product Design Engineering	Engineering Doctorate	N/A	N/A
P2	26	BEng Mechanical engineering / Meng Mechatronics and automation	Engineering Doctorate	2	U of S
P3	33	BEng Industrial design engineering / MEng Manufacturing engineering	Engineering Doctorate	2	U of S
P4	26	MEng Mech & Elec	Engineering Doctorate (Advanced Manufacturing)	3	U of S
P5	25	Mechanical Engineering with Aeronautics	Engineering Doctorate	3	U of S
P6	N/A	N/A	Engineering Doctorate	2	U of S

Table A3-2 – Demographic data for Sim-P2

ID	Age	UG	PG	Current course / occupation	Current year	Institution
P1	35	UG Mechanical Engineering	MSc Global Innovation Management	Research Assistant	N/A	U of S
P2	24	MEng PDE	Integrated Master's	Doctoral student	2	U of S
P3	23	BEng Product Design Engineering	--------------------	CAD Operator/Product Designer	N/A	Industry
P4	23	BA Psychology Hons	--------------------	Doctoral student	2	U of S
P5	22	MEng PDE	Integrated Master's	Undergraduate student	5	U of S
P6	27	Product Design (BSc)	Advanced Manufacturing (MSc)	Doctoral student	4	U of S
P7	26	BSc Mechanical engineering, Uni of Iceland,	MSc Mechatronics and automation, Uni of Strathclyde	Doctoral student	2	U of S
P8	27	MEng PDE	Integrated Master's	Doctoral student	3	U of S
P9	26	BEng Product Design Engineering	--------------------	Doctoral student	4	U of S
P10	26	MEng Sports Engineering	Integrated Master's	N/A (finished phd)	N/A	U of S
P11	31	International Product Design BA (Hons)	Product Engineering Design MSc	Doctoral student	2	U of S

Table A3-3 - Demographic data for Combo-P1

RandID	Age	Gender	UG	PG	Current	Year of study
P1		M	International product design	MSc Product Engineering Design	Doctoral student	2
P2		M	Product Design	MSc Adv. Manufacturing	Doctoral student	3
P3		M	Product Design Engineering MEng	Integrated	Doctoral student	3
P4		M	Product Design and Innovation	---------------------	Doctoral student	1
P5		M	Product Design Engineering MEng	Integrated	Doctoral student	N/A
P6		F	Sports Engineering MEng	Integrated	Research Assistant	

Table A3-4 – Demographic data for Combo-P3

RandID	Age	Gender	UG	Current year	Institution
P1		F	PDE	4	U of S
P2		M	PDE	2	U of S
P3		F	PDE	3	U of S
P4		M	PDE	5	U of S
P5			PDE		U of S
P6		M	PDE	3	U of S
P7		F	PDE	2	U of S
P8		M	PDE	5	U of S
P9		M	PDE	3	U of S
P10		F	PDE	2	U of S
P11		F	PDE	5	U of S
P12		M	PDE	3	U of S
P13		F	PDE	5	U of S
P14		M	PDE	2	U of S
P15		M	PDE	5	U of S
P16		M	PDE	?	U of S
P17		F	PDE	2	U of S
P18		F	PDE	5	U of S
P19		F	PDE	2	U of S
P20		F	PDE	5	U of S
P21		M	PDE	5	U of S
P22		F	PDE	3	U of S
P23		M	PDE	2	U of S
P24		F	PDE	5	U of S
P25		F	PDE	5	U of S
P26		F	PDE	3	U of S
P27		M	PDE	4	U of S
P28		M	PDE	2	U of S
P29		M	PDE	5	U of S
P30		M	MSC…	N/A	U of S

Appendix 3B – Behavioural study

The stimuli creation process (Chapter 6) involved the creation of pairs of design concepts that varied in relative similarity. The design concept used to create those pairs were created in a concept generation experiment to which the researcher made a significant contribution. This study, termed the 'behavioural study', was conducted in parallel with the experiments reported in this thesis, and consisted of a series of cognitive ability tests and design concept generation tasks. The design concept generation tasks involved generating up to three design concepts in response to one of ten open-ended, problem-oriented design tasks within 6 minutes. The design concepts produced in behavioural study were sampled at various stages during the data collection process and used as inputs into three rounds of stimuli creation. The procedure for this experiment was as follows.

Participants. The participants who took part in the study were taken from the 'PDE competent' population described in Section 5.1.1, but these participants were reimbursed £30 for completion of the study. The study was approved by the Department of Design, Manufacturing and Engineering Management Ethics Committee.

Materials. For the design concept generation part of the behavioural study, the participants were given variants of the general materials described in Section 5.1.3. They sketched their concepts on the sheet shown in Figure 5-2 that contained designated spaces for (i) sketches and annotations, and (ii) a description of the concept.

The participants responded to ten design tasks listed in Table A3-5. These tasks are problem-oriented (Sosa, 2018), open-ended ideation tasks. These are tasks that introduce a context and a problem or challenge to be addressed. In contrast, 'solution-oriented' tasks are tasks that specify a desired output, such as a function or category of product that must be created. The tasks were defined based on examples and inspiration from student projects in DMEM and worldwide design competitions. Each task is presented in two sentences, the first being a context and problem to be addressed with examples of the nature of common problems, followed by a short task instruction.

Table A3-5 – Design tasks used in the behavioural study

Task #	Description
DT03	Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.
DT06	Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.
DT07	Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.
DT08	Air travel may be problematic for wheelchair users due to difficulties with manual handling and manoeuvring. Generate concepts for products that may make airports and planes more wheelchair-friendly.
DT09	Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.
DT12	Chores such as cooking and cleaning may be difficult for wheelchair users due to space and height limitations. Generate concepts for products that may facilitate domestic chores for wheelchair users.
DT14	Train stations and airports are often congested due to many people transporting large items in a confined space. Generate concepts for products that may reduce congestion in transportation hubs.
DT15	Rain and wind make it difficult for pedestrians to keep dry and pose dangers e.g. slipping; falling trees. Generate concepts for products to reduce the discomfort and danger of poor weather for pedestrians.
DT19	Inner city pavements are often congested due to large numbers of people and obstructions like street furniture; parked cars; other pedestrians. Generate concepts for products that may reduce congestion for inner city pedestrians.
DT20	Dog excrement on pavements is unsightly and unhygienic but its disposal may be unpleasant and unhygienic for dog owners. Generate concepts for products that may improve dog excrement disposal for dog owners.

Procedure. In the design concept generation portion of the experiment, participants responded to ten design briefs by creating up to three concepts within a time limit of 6 minutes. The instructions given to the participants were:

> *"During this part of the experiment, you will be asked to generate concepts in response to 10 different design problems which will be shown to you on a computer screen. The tasks, and the concepts you generate during the tasks, are intended to represent the conceptual,* creative design phase*.*

For each design problem, you will be given 6 minutes to generate and sketch up to three concepts. Please try to generate one concept at a time*, and only once the first concept has been generated and sketched should you proceed to generate the next concept."*

Participants were also informed that the sketches of their concepts did not need to be highly detailed, but that they should try to convey the key components of the concept.

Outputs. An example of the kinds of concept created is shown in Figure A3-1. In this example the participant has created a sketch, annotations, directional arrows indicating movement of the object, and arrows linking annotations to parts of the sketch. They have also provided a description of the concept in the bottom-right of the image which could have alternatively been placed in the grey box that prompts the participant for a description of the product. In this box they have labelled the concept 'waste-compactor'.

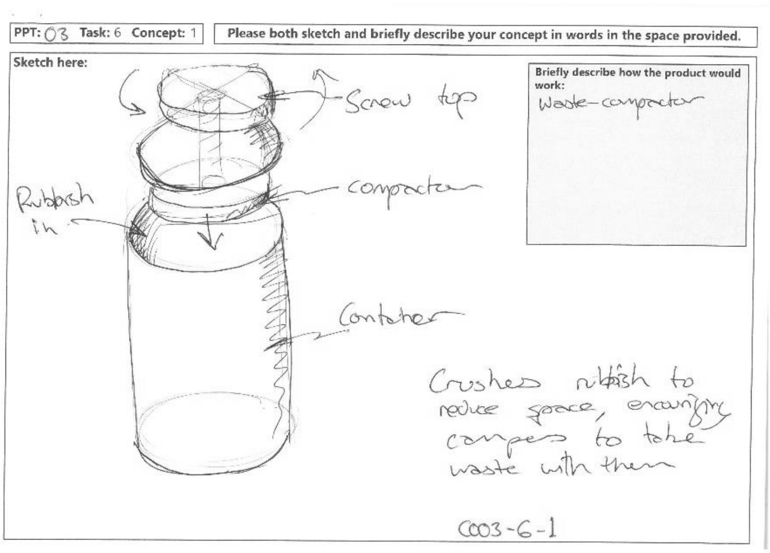


Figure A3-1 – An example of a design concept produced by a participant in the concept generation task

APPENDIX 4 – STIMULI CREATION

APPENDIX 4A – AN EXAMPLE OF A CODED CONCEPT USING THE FBS SCHEME

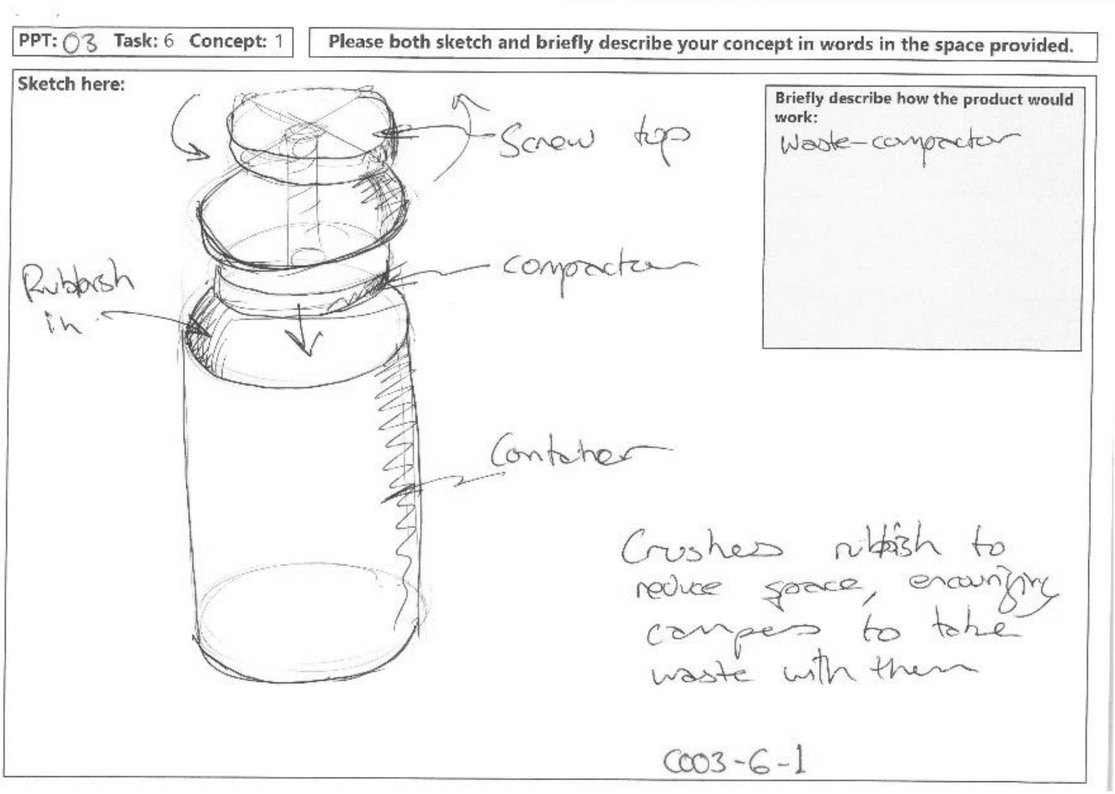


Figure A4-1 – An example of a design concept used in the creation of the stimuli pairs.

Table A4-1 – Codes applied to the design concept shown in Figure A4-1

Info.	Concept	C003-6-1
	Design task	Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.
	Goal of the problem	Reduce the negative environmental impacts of camping
Codes	F	Facilitate storage of waste during camping trips
	B	Compact rubbish
		Contains rubbish
	S	Rubbish receptacle
		Mechanical compacting device

APPENDIX 4B – EXAMPLES OF PAIRS AT THREE LEVELS OF ONTOLOGICAL DISTANCE

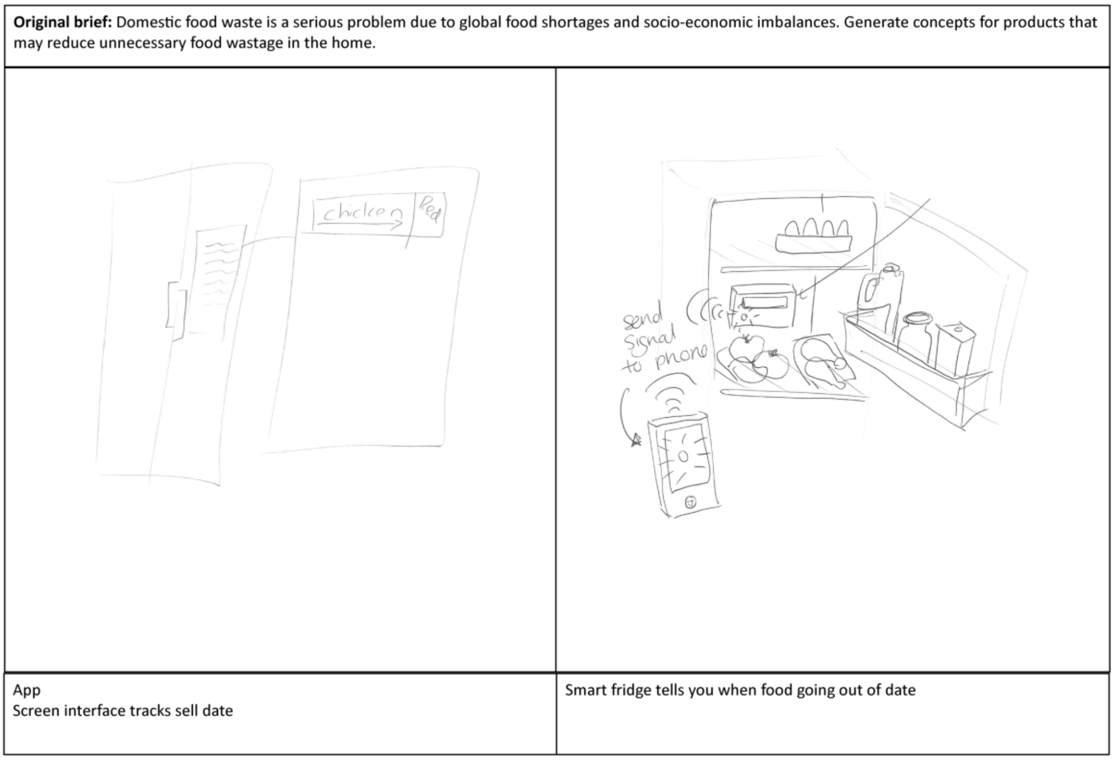


Figure A4-2 - A visual example of a pair of design concepts at ontological distance 1

Table A4-2 - An example of the purpose, function and means of a pair of design concepts at ontological distance 1

	Concept (left)	Concept (right)	Category location
Purpose (brief)	Reduce unnecessary food waste in the home	Reduce unnecessary food waste in the home	Common (by default)
Purpose (derived)	"reduce or eliminate instances of food expiring before it is used".	"reduce or eliminate instances of food expiring before it is used".	Common
Purpose (derived)	"make the user aware of the expiration-state of the food".	"make the user aware of the expiration-state of the food".	Common
Function	Display information about the upcoming expiry date of food, at a given time such that a user would have a chance to act upon the information that is displayed.	Display information about the upcoming expiry date of food, at a given time such that a user would have a chance to act upon the information that is displayed.	Common
Means	A 'smart' refrigerator (with external display screen on door).	A 'smart' refrigerator (that communicates with smartphone).	Common

The top-level purpose is specified in the brief. Pairs were only created using design concepts from the same original brief and so there is always a minimum degree of commonality by default.

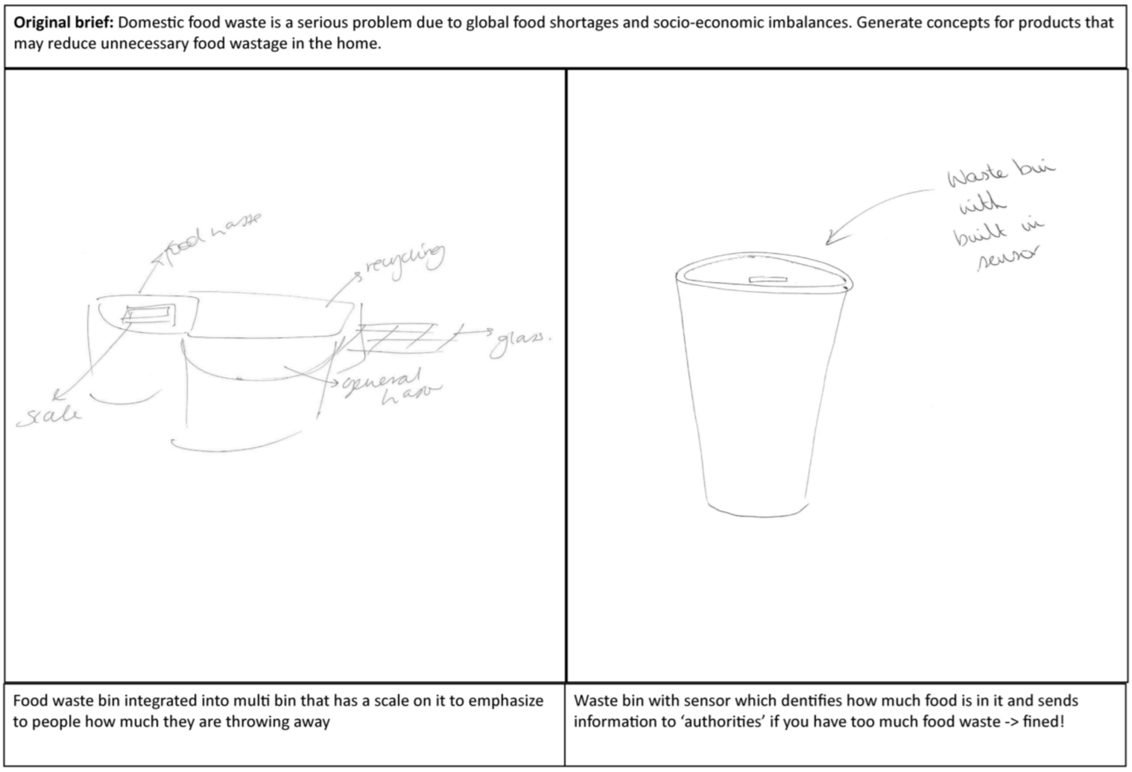


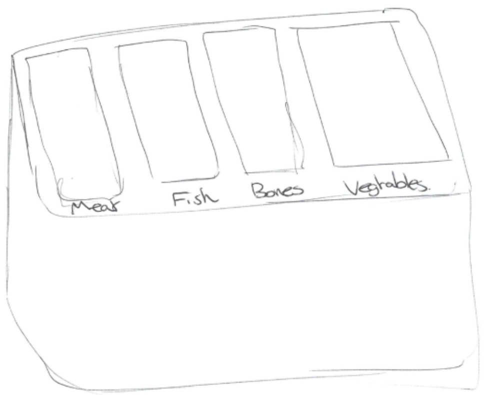
Figure A4-3 - A visual example of a pair of design concepts at ontological distance 3

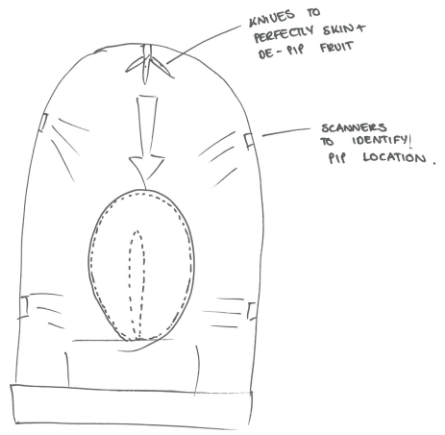
Table A4-3 - An example of the purpose, function and means of a pair of design concepts at ontological distance 3

	Concept (left)	**Concept (right)**	**Category location**
Purpose (brief)	Reduce unnecessary food waste in the home	Reduce unnecessary food waste in the home	Common (by default)
Purpose (derived)	Discourage people from producing excessive food waste	Discourage people from producing excessive food waste	Common
Purpose (derived)	Discourage people from producing excessive food waste via behavioural intervention	Discourage people from producing excessive food waste via behavioural intervention	Common
Function	Detect and display quantity of food waste	Detect quantity of food waste and send information to authorities at pre-defined limit	Different
Means	Food waste bin with sensor and display	Food waste bin with sensor and external communication with authorities	Different

The top-level purpose is specified in the brief. Pairs were only created using design concepts from the same original brief and so there is always a minimum degree of commonality by default.

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.





Differen sections of the bins so that the meat can be given to animals / taken to zoo, fish to cats or sharks, bones turned into glue and vegetables compost	Perfect peeler Reduced fruit / veg cost in peeling process

Figure A4-4 - A visual example of a pair of design concepts at ontological distance 5

Table A4-4 - An example of the purpose, function and means of a pair of design concepts at ontological distance 5

	Concept (left)	Concept (right)	Category location
Purpose (brief)	Reduce unnecessary food waste in the home	Reduce unnecessary food waste in the home	Common (by default)
Purpose (derived)	Reprocess or repurpose waste	Prevent food waste during preparation and consumption	Different
Purpose (derived)	Facilitate the transformation of waste into animal food	Prevent food waste caused by inefficient peeling of fruit or vegetables	Different
Function	Store waste	Peel fruit or vegetables (with low waste of edible material)	Different
Means	Four-compartment bin	Bell-jar shaped peeling system	Different

The top-level purpose is specified in the brief. Pairs were only created using design concepts from the same original brief and so there is always a minimum degree of commonality by default.

APPENDIX 4C – STIMULI SET B

Table A4-5 – Stimuli Set B

DT	Align	Unique ID	Also in	DT	Align	Unique ID	Also in
	1	Sb_C01	Sc		1	Sb_C21	Sc
	1	Sb_C02	Sc		1	Sb_C22	Sc
	2	Sb_C03	-		2	Sb_C23	Sc
	2	Sb_C04	-		2	Sb_C24	-
	3	Sb_C05	-		3	Sb_C25	Sc
1	3	Sb_C06	Sc	3	3	Sb_C26	-
	4	Sb_C07	Sc		4	Sb_C27	-
	4	Sb_C08	-		4	Sb_C2	Sc
	5	Sb_C09	-		5	Sb_C29	Sc
	5	Sb_C10	Sc		5	Sb_C30	Sc
	1	Sb_C11	Sc		1	Sb_C31	Sc-W
	1	Sb_C12	Sc		1	Sb_C32	Sc-W
	2	Sb_C13	Sc		2	Sb_C33	Sc-W
	2	Sb_C14	Sc		2	Sb_C34	Sc-W
	3	Sb_C15	Sc		3	Sb_C35	Sc-W
2	3	Sb_C16	-	4	3	Sb_C36	Sc-W
	4	Sb_C17	-		4	Sb_C37	Sc-W
	4	Sb_C18	Sc		4	Sb_C38	Sc-W
	5	Sb_C19	-		5	Sb_C39	Sc-W
	5	Sb_C20	-		5	Sb_C40	Sc-W

Note. *DT = design task, Align = alignability level from categorisation, 'Also in' indicates that this pair also appear in Stimuli Set C (Sc), or the warmup booklets associated with Stimuli Set C (Sc-W).*

Sb_C01

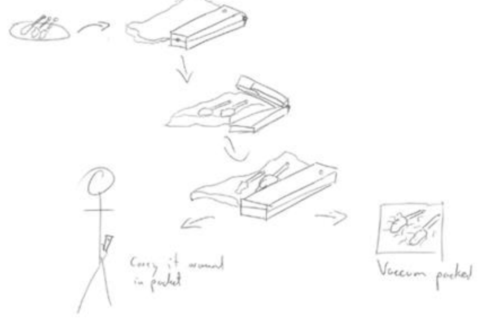
Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

App
Screen interface tracks sell date

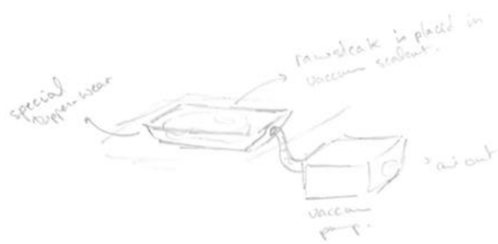
Smart fridge tells you when food going out of date

Sb_C02

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



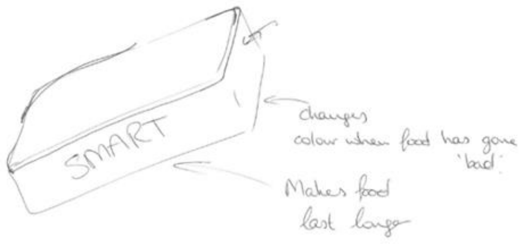
Pocket sized vacuum packer



Easy pack! Many food items have long expirey dates on them but should be consumed wtihin few days if the packet is opened. A device that could repack food would solve this issue.

Sb_C03

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

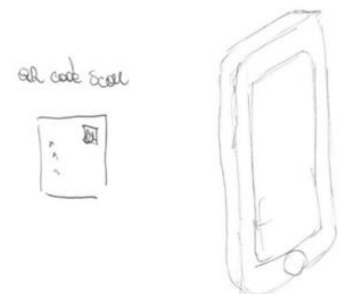


People throw away leftovers when they have gone 'bad'.
'Smart box' changes colour when it is unsafe to eat food and new airtight design keeps food fresh longer

Labels react with chemicals and bacteria inside food to show if it is expiring

Sb_C04

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

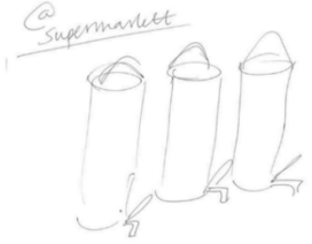


smartphone alert on the expired date (some days before)

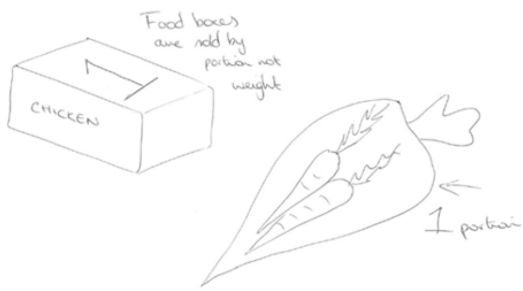
a scanner that saves expiry dates of products you put into the fridge and notifies you when it's closing by so you consume it

Sb_C05

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



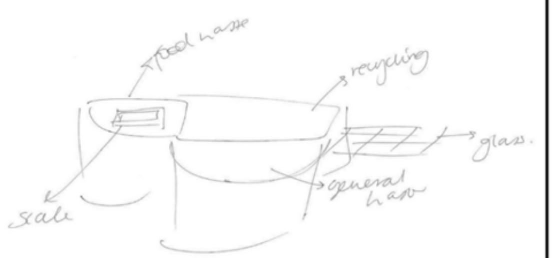
All food in supermarket can be bought in exact quantities
Less food waste as someone dining alone can buy less than a family -> less packaging -> could bring own container to fill up



Food is sold per portion not by weight
People will cook right amount without having too much

Sb_C06

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

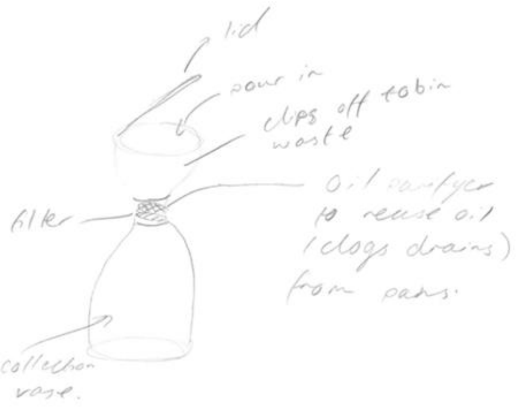


Food waste bin integrated into multi bin that has a scale on it to emphasize to people how much they are throwing away

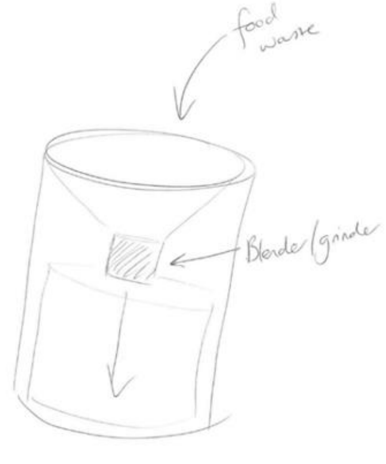
Waste bin with sensor which dentifies how much food is in it and sends information to 'authorities' if you have too much food waste -> fined!

Sb_C07

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



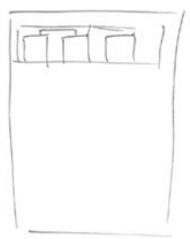
Oil is expensive and clogs drains. A filter means it could be reused



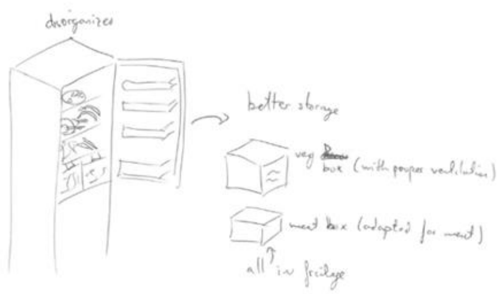
With the fod being grinded and blended it takes up less space and it is ready for recycling. People always like to have a gadget in their house too... even if its usefulness is questionable

Sb_C08

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



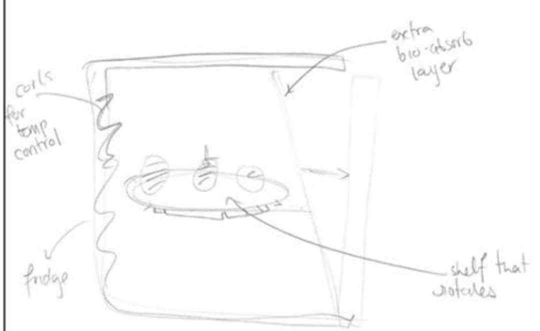
A shelf that hides the more in-date food and presents ones about to go off

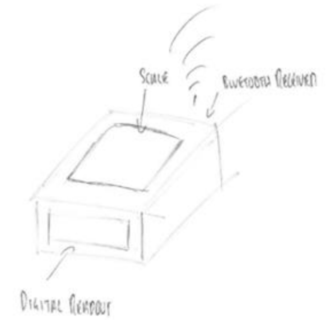


Proper organised storage for food

Sb_C09

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

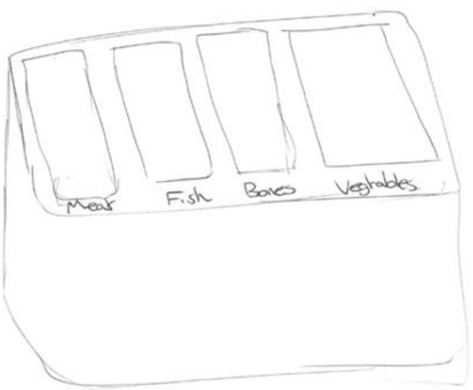


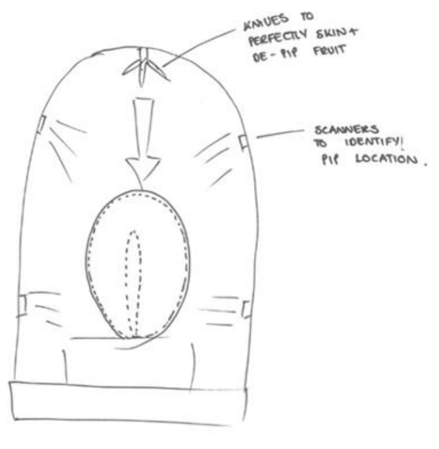


Reduce food spoilage via fridge door mechanism that has extra bio absorbant layer and rotates food away from opening towards cooling coils.	Digital smart scale syncs with phone on media deice to show not only appropriate mass but visually what components of food should be removed and where to find them

Sb_C10

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

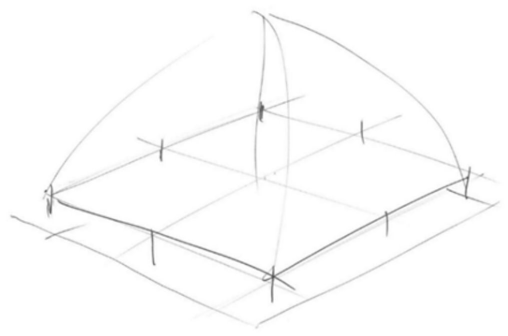




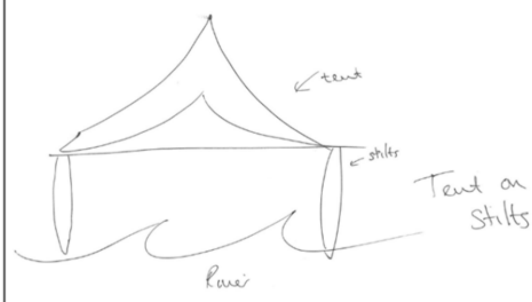
Differen sections of the bins so that the meat can be given to animals / taken to zoo, fish to cats or sharks, bones turned into glue and vegetables compost	Perfect peeler Reduced fruit / veg cost in peeling process

Sb_C11

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Elevated tent

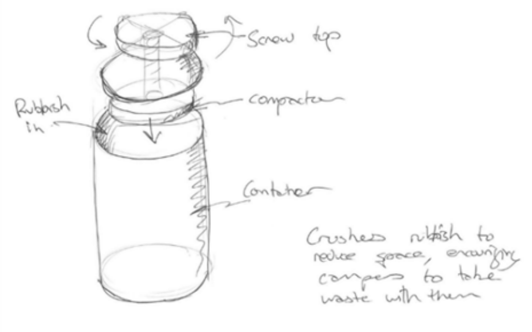


Tent on stilts... can camp anywhere and can avoid staying in protected areas

Sb_C12

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

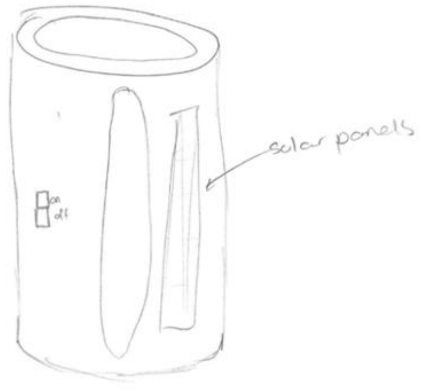
Box filled with rubbish and then squeezed to make rubbish take up as little space as possible. Detachable at bottom to then put in bag



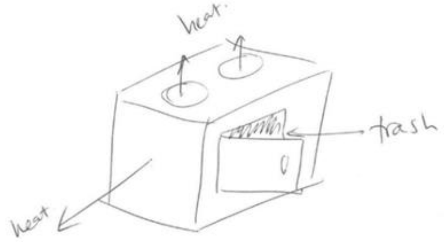
Waste compactor

Sb_C13

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



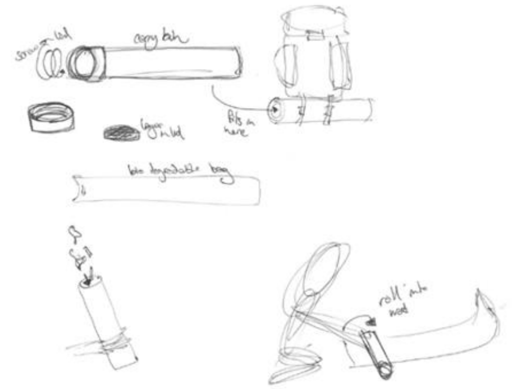
A trash can that charges up throughout the day and can then burn the trash and it becomes and stove



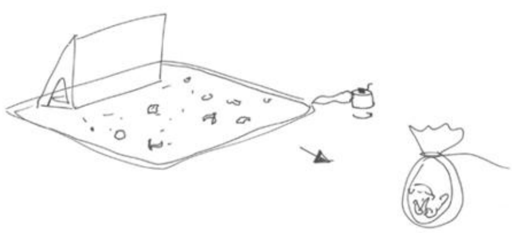
Portable incinerator. The rubbish is burned, heat is recycled through "fireplace" on side or "hob" on top for cooking foods. Each of the two is switched between. Stops rubbish being left behind.

Sb_C14

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Leaving rubbish behind when packing up is a common issue. This is an air tight cyclidrical bin that fits in rolled up with your sleeping mat which is attached to bag using standard straps they already have. Non smelling layer with disposal at your next stop

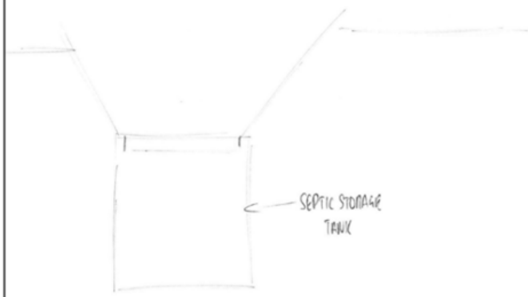


Mat that gets put down in camping site on top of grass. When campers gone the mat scrunches up into a bag. Catching rubbish to be thrown away later.

Sb_C15

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

Bin that sits outside tent wall, but can be opened from inside

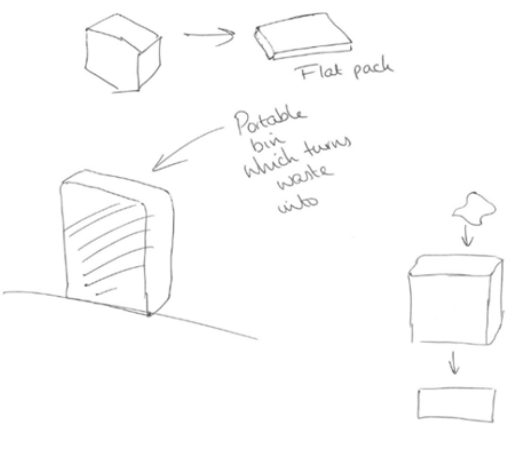


Septic storage tank to prevent pollution to nutrient rich mud and potential insects. Portable and sealable after use with biodegradeable funnel for ease of storage.

Sb_C16

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

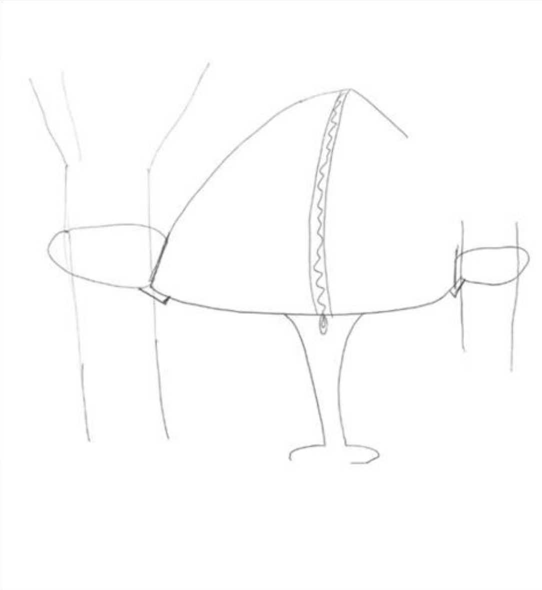
Food plates and cups will be disolved by environment so the materials have to recyclable



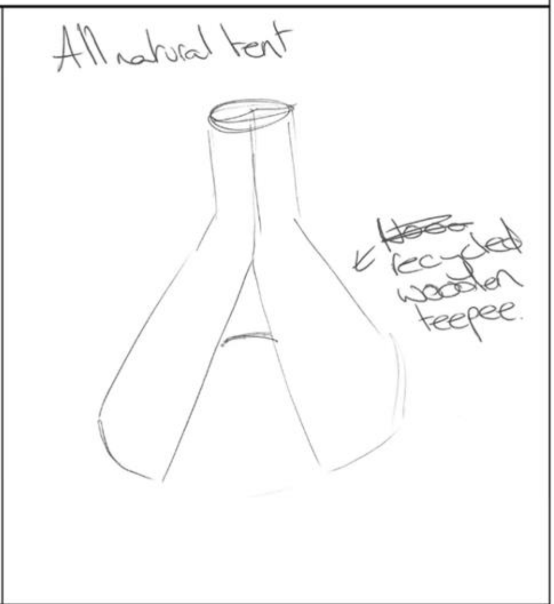
Bin where your waste is transformed into bio-degradeable objects. Reduces litter

Sb_C17

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



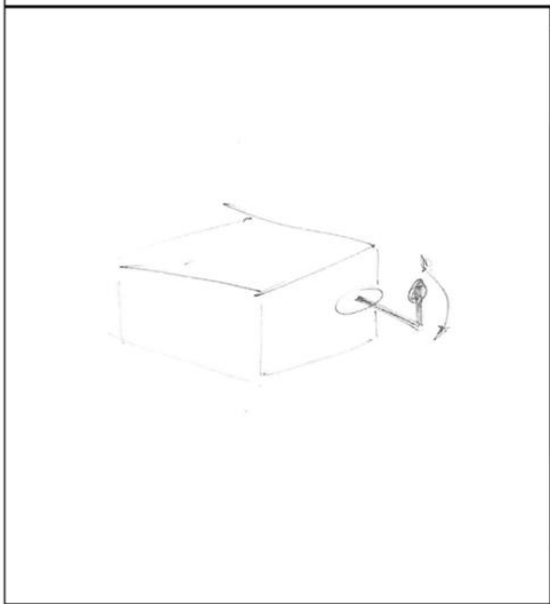
Hammock tents where a tent is suspended between two trees and is supported by one middle pole to prevent the grass and ground area being ruined



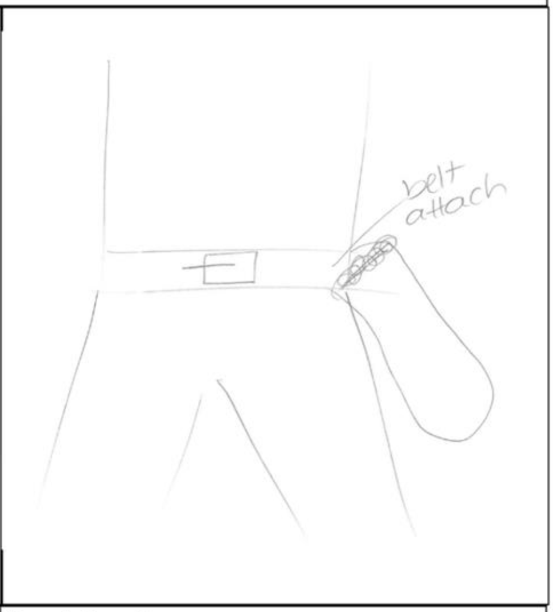
it can stay in the landscape and will biodegrade

Sb_C18

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Users place food waste in bin. It is then hand ground and users can spread on vegitation before leaving

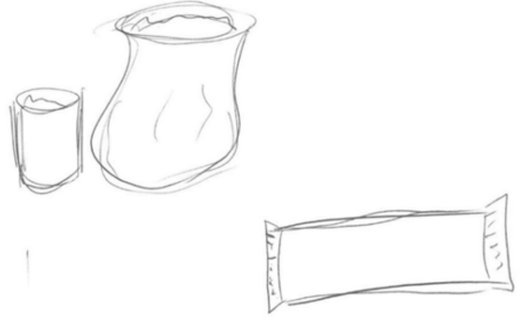


Personal refuse sack

Sb_C19

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

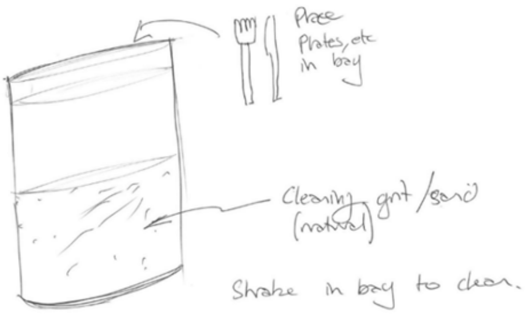
Tent elevated off ground - less damage to plants on ground



Packaging redesign. Edible, biodegradeable or suitable for throwing on the camp fire - reduces litter

Sb_C20

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

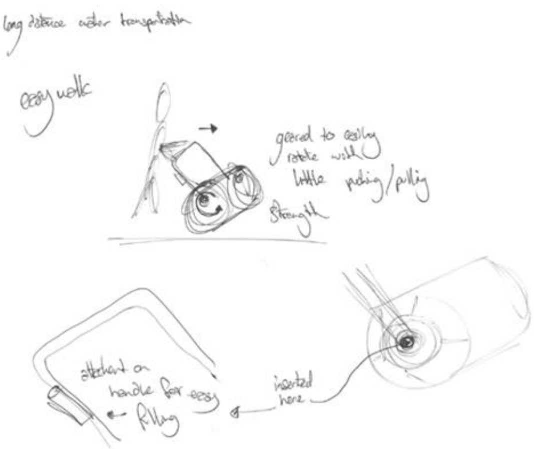


Shake to clean - no soaps!

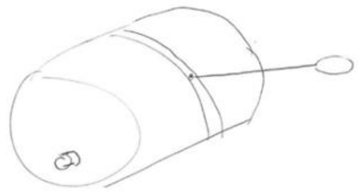
'native noises' - birds, wind, leaves, water etc.

Sb_C21

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



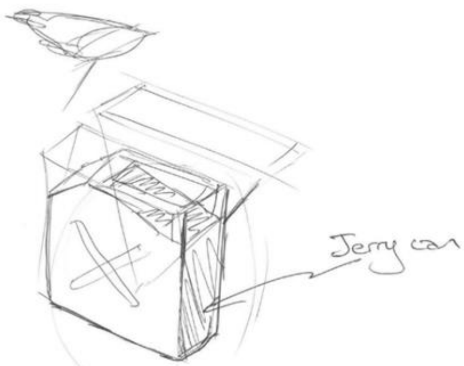
Water barrel with attached handle for pushing water over long distances. For those who need to walk to collect water



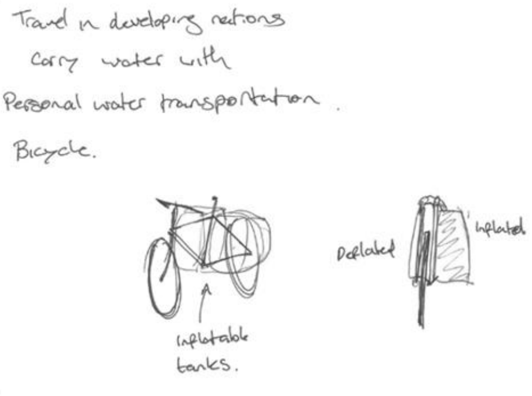
Barrell with string attached to ball-bearing ring that can be easily dragged / rolled when full

Sb_C22

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Jerry can panier for bicycles

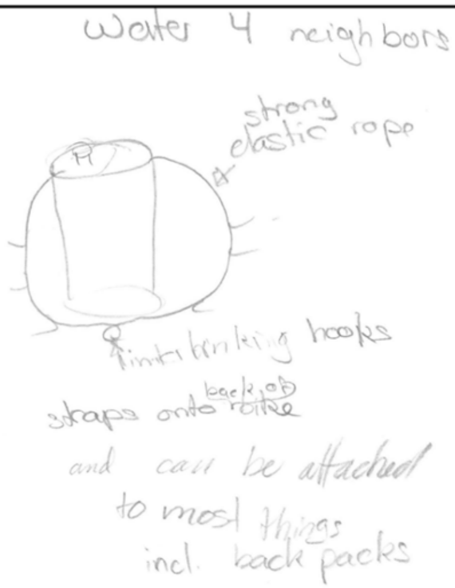


Tank either side of rear wheel.
Hooks on to conventional bicycle

Sb_C23

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

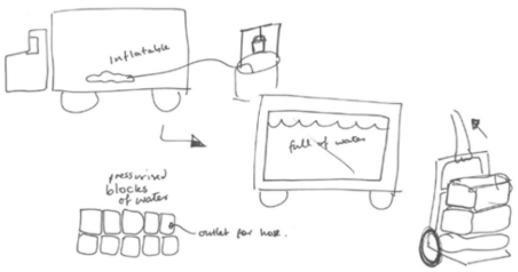
Water container bag, to make it easier to carry water



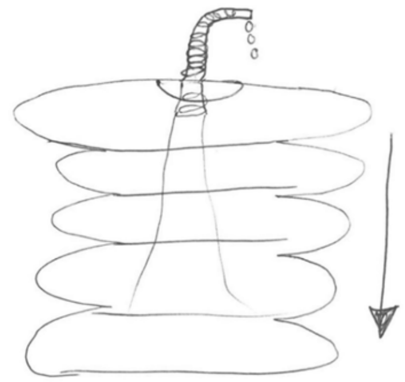
Can be transported / used by anyone

Sb_C24

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Felxible 'bags' that can be filled up. No spillage, can stack together. Needs wheeled transpotation -e.g. suitcase, vehicle



Large foldable bottles could be sent that people would fill up and the water would then be filtered cleanly through a pump

Sb_C25

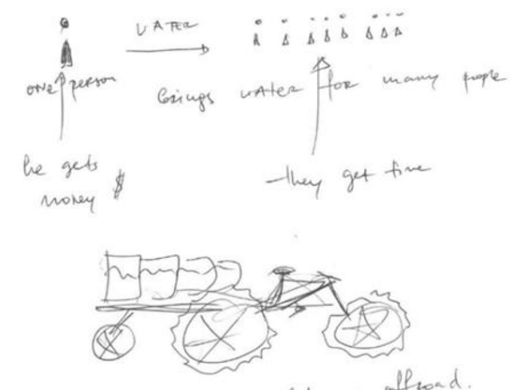
Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

Bus-mounted water transport

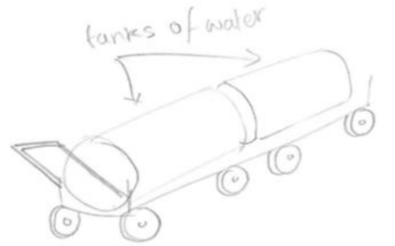
Heavy lift drones for mass delivery

Sb_C26

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



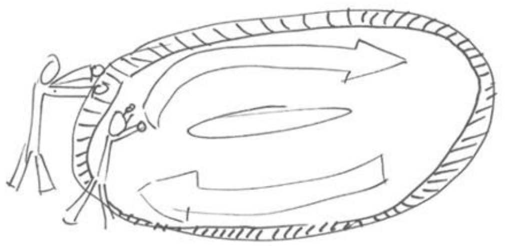
cargo bike



Using a pull bar and a lot of tough thick tyre treads for long rough terrain it could be easy for 1/2 people to transport a large amount of water

Sb_C27

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

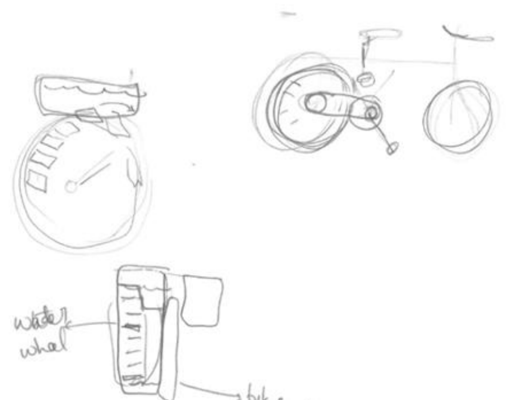


A large water balloon that is a hoop so when pushed, it travels like a tank track. Thin tank tracks are on the outside to protect innards from rough terrain

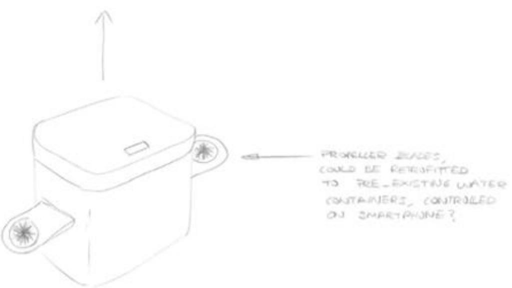
small solar/electric powered vehicles for transportation / job creation

Sb_C28

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



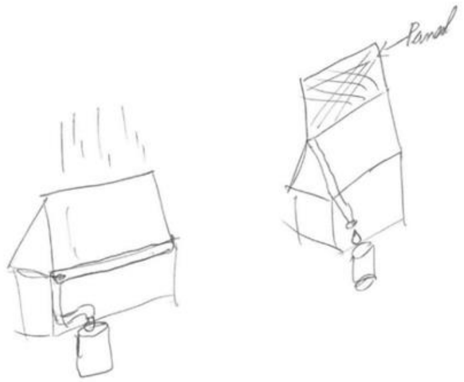
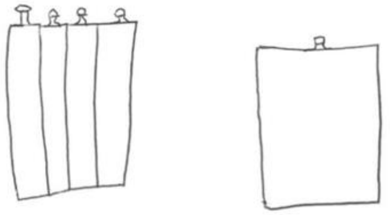
Bike that uses H20 for conservation of momentum (like a water wheel) - once bike is in motion it can move a long ways w/ less effort



Solar powered water drones that are cheap and easy to look after, clean, repair, robust stucture

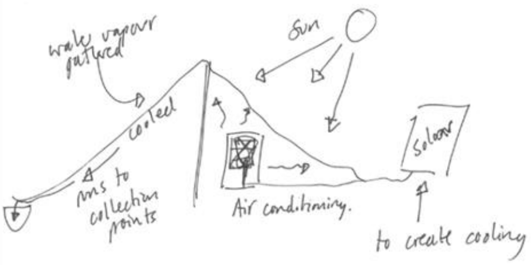
Sb_C29

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

	
Collection of rain water Panel that retains moisture to collect it	Individual water bottles are all rectangular shaped to be easily transported in a box, lots can be stored together

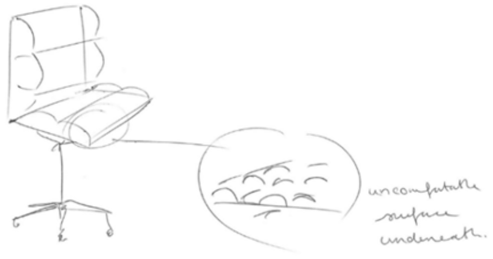
Sb_C30

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

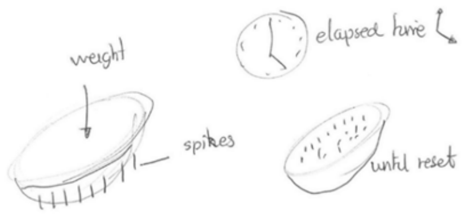
	
Designed to be filled and used over varying conditions. Obviously consideration must be given to the max. weight people can actually push, relative to the capacity of the holder	No transportation Made at source from water in atmosphere

Sb_C31

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.



Inflatable desk chair cushions which deflate as you use them making chair more and more uncomfortable until you get up



Over time the weight of the person initiates a compression of the chair to which spikes come out and only resets after 15 minute intervals

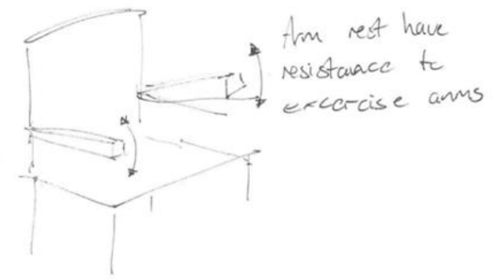
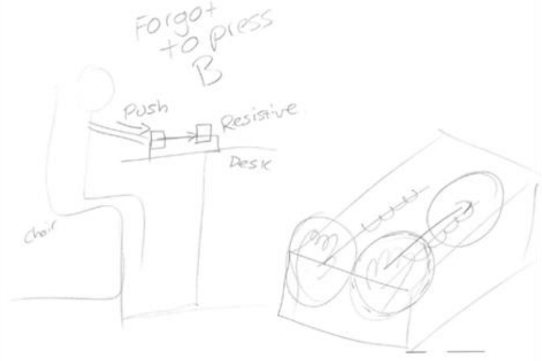
Sb_C32

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.

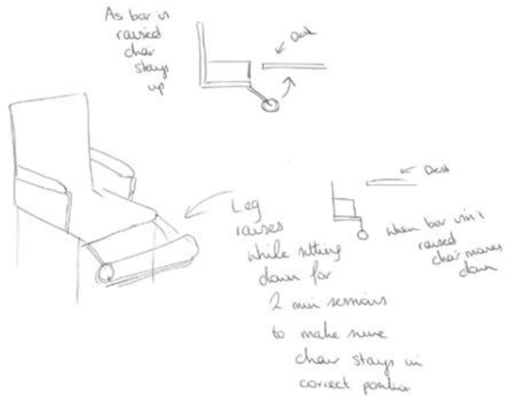
Electrical yoga balls that change where your balance is to make you tense to stay upright

Similar to using a yoga ball to sit on. This chair has a back which is recline-able and moves to the position of the user. This prevents sore backs as the ball chair can flex to any sitting position.

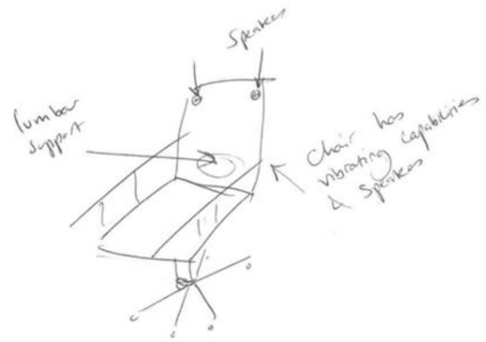
Sb_C33

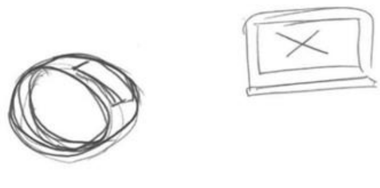
Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.	
	
User can exercise arms by using arm rest	Resistive exercise machine for hands and arms

Sb_C34

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.	
	
Pedal while seated / working	Weight chair, incentive to keep seating position

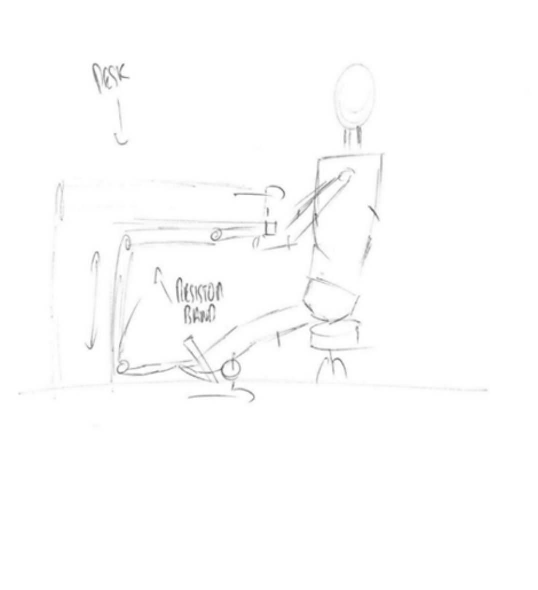
Sb_C35

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.	
	
Alarm to alert you to do the exercise. Flashing light for glute tensing exercises: when light is on - tense glutes Could get harder the longer they use it	Chair vibrates to remind you to exercise. Pre-recorded message plays through speakers or headphones, instructing you of exercises to complete.

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.	
	
Seat w/pressure sensors, if you have been sitting to long wifi is blocked until you get up and walk or stand for a few min.	Wearables should interfere / lock your laptop unless you do brief exercise. Although this could interfere during deadlines

Sb_C37

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.

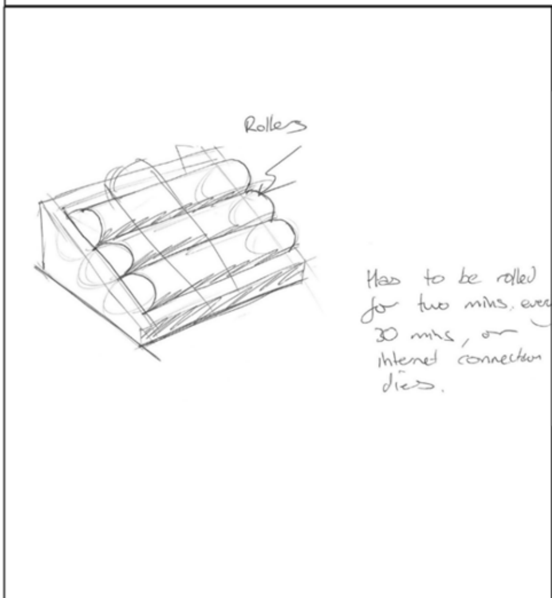


Network of resistor bands attempt to retain tension as user pulls at under mounts at hands / feet using most muscle groups

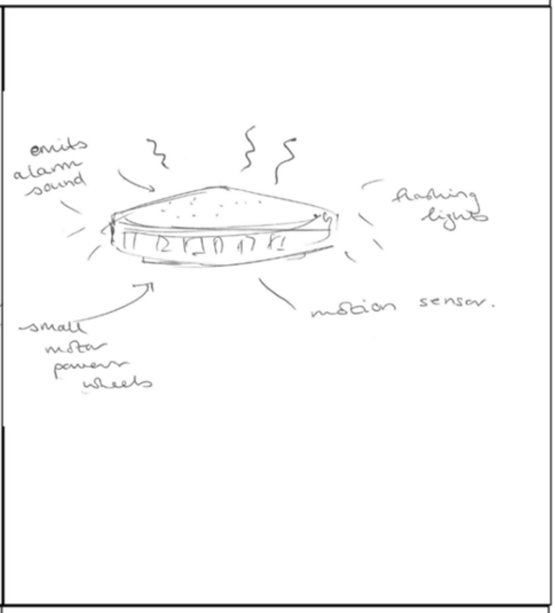
User holds onto handles and engages abs whilst twisting on chair

Sb_C38

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.



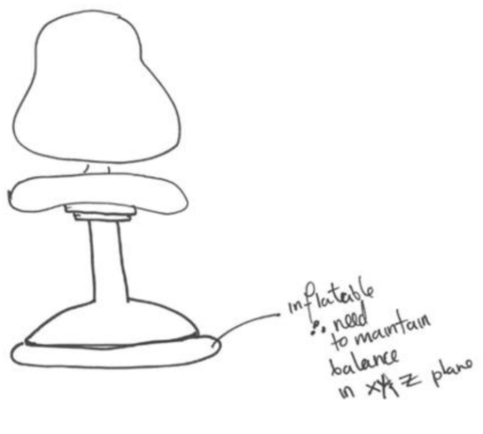
Rollers for feet, connected to computer - no roll, no internet!



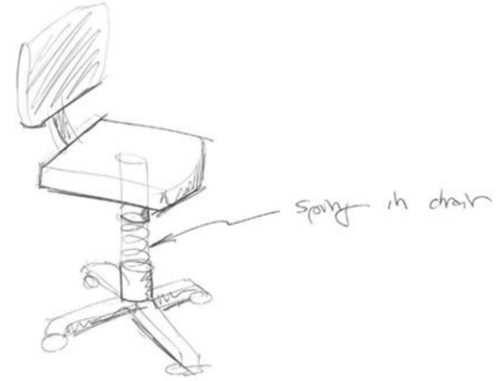
When robot hasn't sensed movement in room for x amount of time it starts going off and tries to run away until you catch it and turn it off

Sb_C39

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.



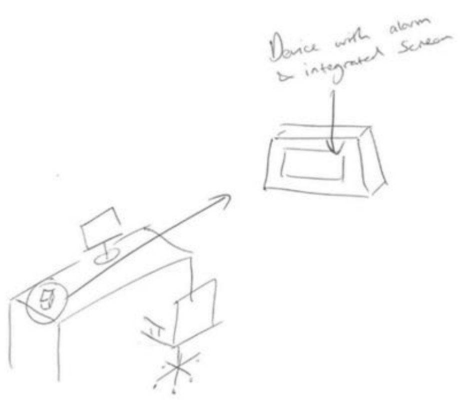
Seat that moves and you need to adjust your position / maintain balance



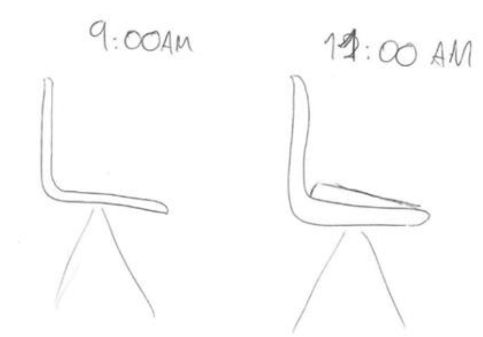
Bounce your way to fitness

Sb_C40

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.



Device alarm goes off periodically to remind you to exercise. Screen shows you the exercises user needs to do and the user replicates what is shown on the screen.



Chair that lifsts the user off the chair slightly over time to incentivise getting up more often

APPENDIX 4D – STIMULI SET C

Table A4-6 – Stimuli set C showing changes between stimuli sets B and C

			Reason for modification	
DT	Unique ID	Also in	A (left concept)	B (right concept)
1	Sc_C01	Sb		
	Sc_C02	Sb		
	Sc_C03	-	Dual functions	
	Sc_C04	-	Lack of clarity	
	Sc_C05	-	Service system elements	Service system elements
	Sc_C06	Sb		
	Sc_C07	Sb		
	Sc_C08	-		Passive
	Sc_C09	-	Dual function	Dual function
	Sc_C10	Sb		
2	Sc_C11	Sb		
	Sc_C12	Sb		
	Sc_C13	Sb		
	Sc_C14	Sb		
	Sc_C15	Sb		
	Sc_C16	-	Passive	
	Sc_C17	-		Passive
	Sc_C18	Sb		
	Sc_C19	-		Passive
	Sc_C20	-		Lack of clarity – traded for equivalent concept
	Sc_C21	Sb		
	Sc_C22	Sb		
	Sc_C23	Sb		
	Sc_C24	-	Service system	
	Sc_C25	Sb		
	Sc_C26	-	Service system – edited sketch	
	Sc_C27	-		Service system
	Sc_C28	Sb		
	Sc_C29	Sb		
	Sc_C30	Sb		

Note. DT = design task, 'Also in' indicates that this pair also appear in Stimuli Set B (Sb)

Sc_C01

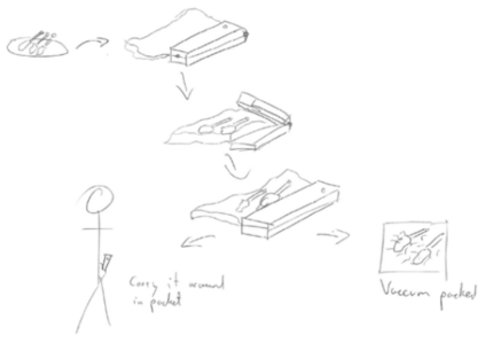
Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

App
Screen interface tracks sell date

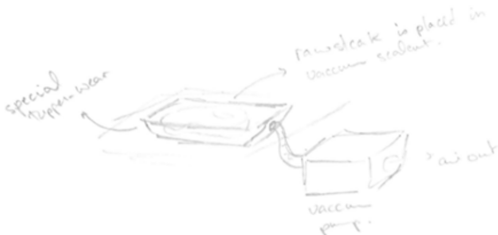
Smart fridge tells you when food going out of date

Sc_C02

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



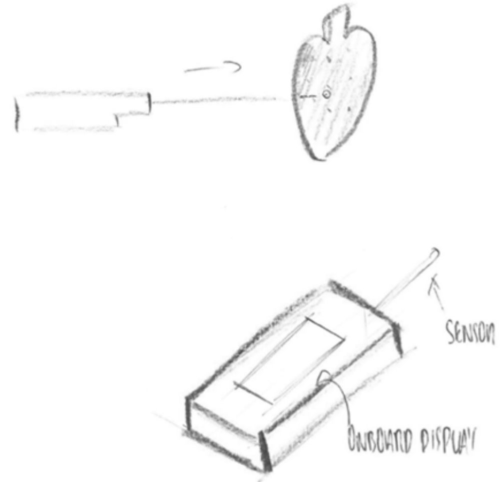
Pocket sized vacuum packer



Easy pack! Many food items have long expiry dates on them but should be consumed wthin few days if the packet is opened. A device that could re-pack food would solve this issue.

Sc_C03

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



Digital sensor to establish bacteria content within food, allowing user to assess accurately safety of food

Labels react with chemicals and bacteria inside food to show if it is expiring

Sc_C04

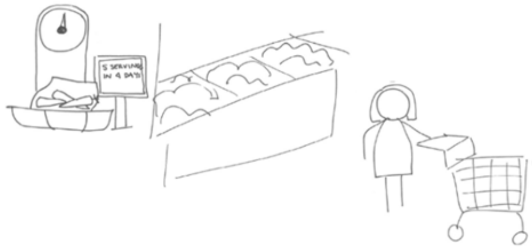
Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

Smart fridge that senses amount left of food and gives examples of recipes

a scanner that saves expiry dates of products you put into the fridge and notifies you when it's closing by so you consume it

Sc_C05

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

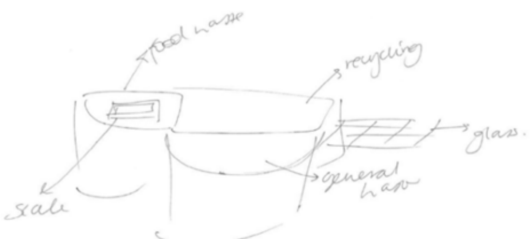


Along with scales in supermarkets for weighing loose veg, it would tell the customer how many servings it contained related to how much time there was

Puts a cap on what you can buy => you won't waste as much

Sc_C06

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

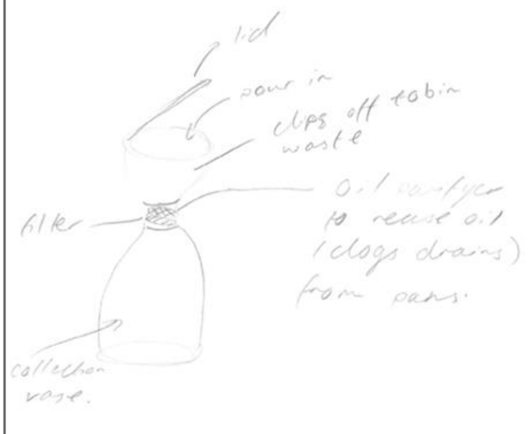


Food waste bin integrated into multi bin that has a scale on it to emphasize to people how much they are throwing away

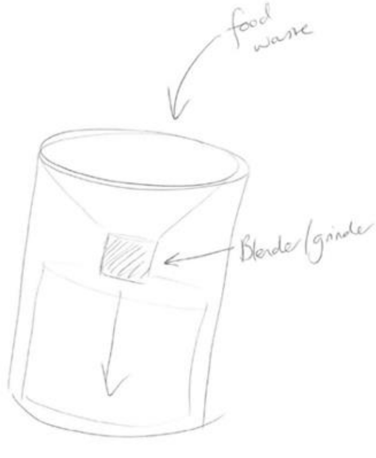
Waste bin with sensor which dentifies how much food is in it and sends information to 'authorities' if you have too much food waste -> fined!

Sc_C07

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



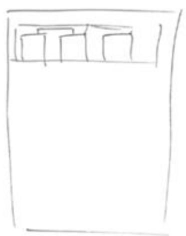
Oil is expensive and clogs drains. A filter means it could be reused



With the fod being grinded and blended it takes up less space and it is ready for recycling. People always like to have a gadget in their house too... even if its usefulness is questionable

Sc_C08

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



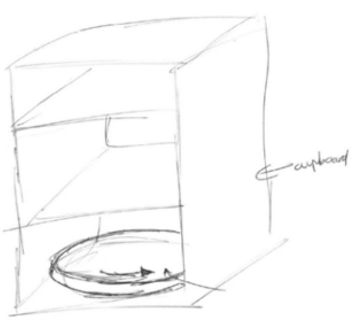
A shelf that hides the more in-date food and presents ones about to go off

Making too much and throwing it away is a huge food waste contributor

Controlling sizes before cooking.

Sc_C09

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.

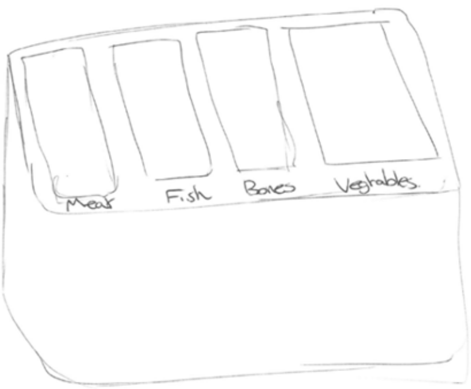


Rotating base rotates at a certain amount each day so that you use the food at the back of the cupboard as well as the stuff at the front

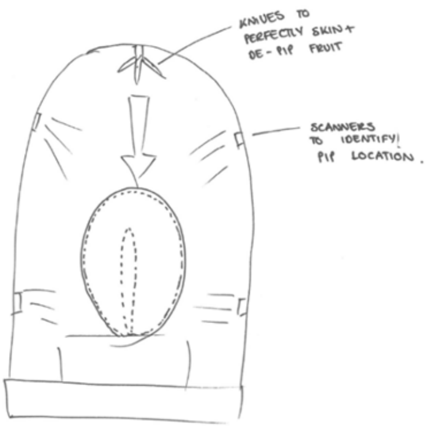
Funnel which counts portion sizes of pasts and grains

Sc_C10

Original brief: Domestic food waste is a serious problem due to global food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home.



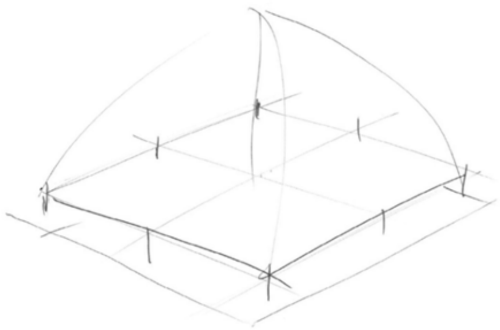
Differen sections of the bins so that the meat can be given to animals / taken to zoo, fish to cats or sharks, bones turned into glue and vegetables compost



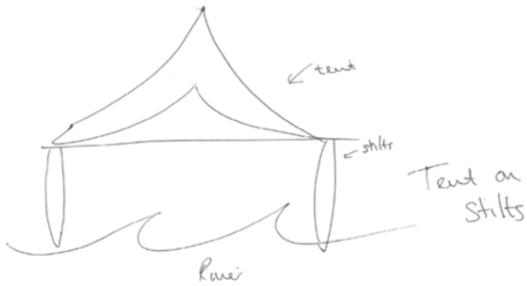
Perfect peeler
Reduced fruit / veg cost in peeling process

Sc_C11

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Elevated tent

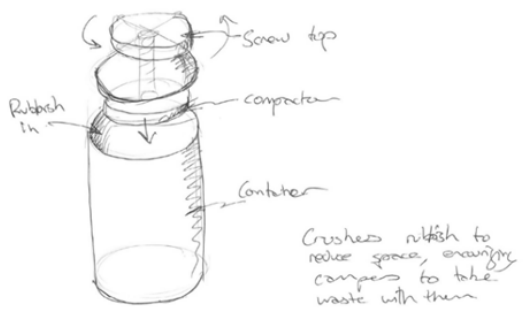


Tent on stilts... can camp anywhere and can avoid staying in protected areas

Sc_C12

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

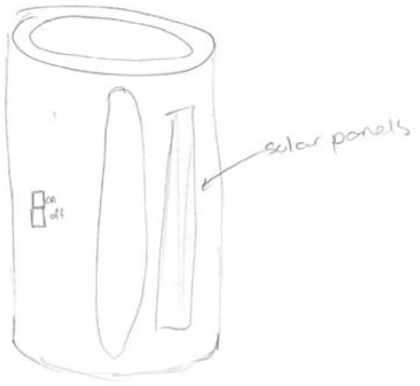
Box filled with rubbish and then squeezed to make rubbish take up as little space as possible. Detachable at bottom to then put in bag



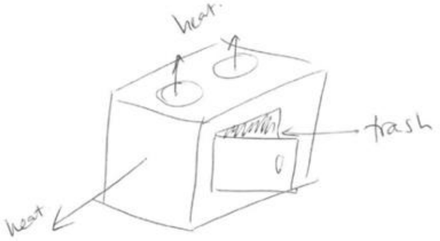
Waste compactor

Sc_C13

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



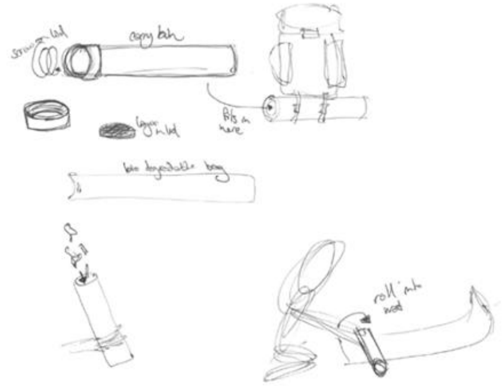
A trash can that charges up throughout the day and can then burn the trash and it becomes and stove



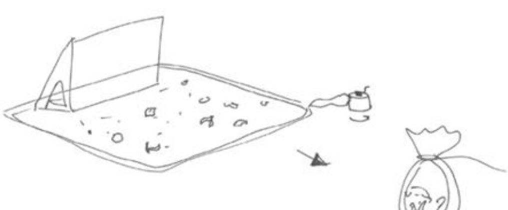
Portable incinerator. The rubbish is burned, heat is recycled through "fireplace" on side or "hob" on top for cooking foods. Each of the two is switched between. Stops rubbish being left behind.

Sc_C14

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Leaving rubbish behind when packing up is a common issue. This is an air tight cyclidrical bin that fits in rolled up with your sleeping mat which is attached to bag using standard straps they already have. Non smelling layer with disposal at your next stop

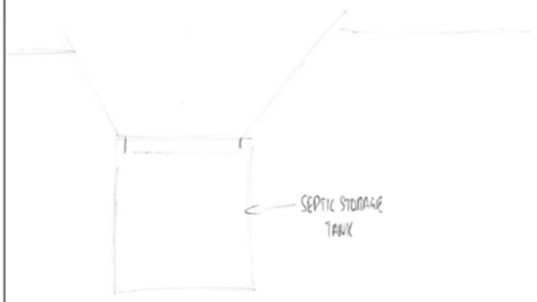


Mat that gets put down in camping site on top of grass. When campers gone the mat scrunches up into a bag. Catching rubbish to be thrown away later.

Sc_C15

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

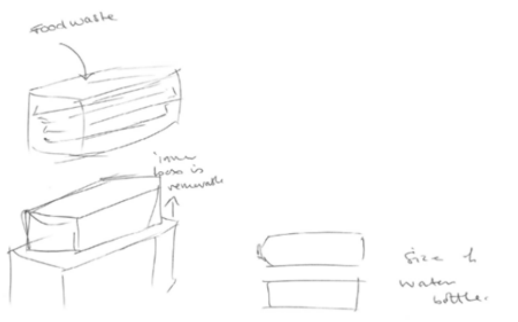
Bin that sits outside tent wall, but can be opened from inside



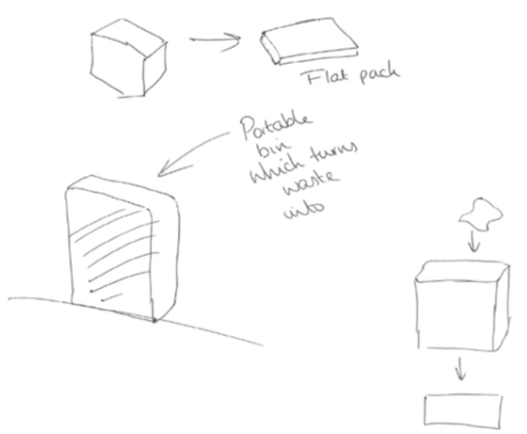
Septic storage tank to prevent pollution to nutrient rich mud and potential insects. Portable and sealable after use with biodegradeable funnel for ease of storage.

Sc_C16

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



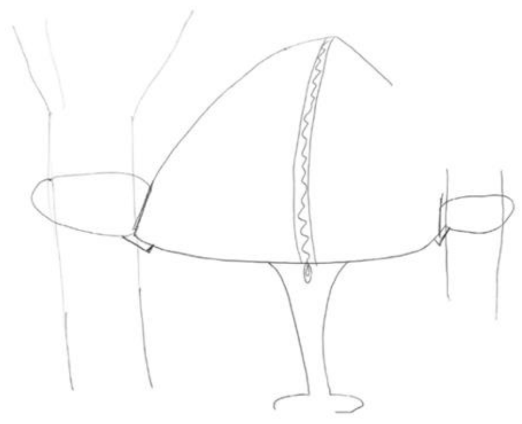
Uses organic chemicals to speed up composting process so food waste can be composted overnight then left at campsite



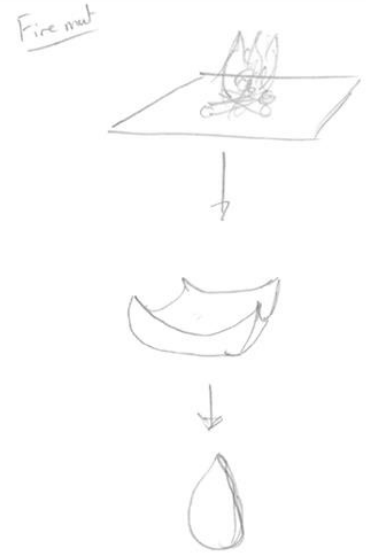
Bin where your waste is transformed into bio-degradeable objects. Reduces litter

Sc_C17

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



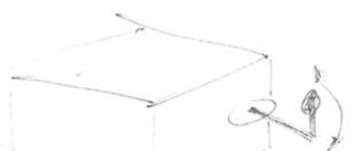
Hammock tents where a tent is suspended between two trees and is supported by one middle pole to prevent the grass and ground area being ruined



A fire proof mat that will protect the ground below so a fire can be started without scarring the ground, the ash could then be taken away to be suitably recycled

Sc_C18

Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.

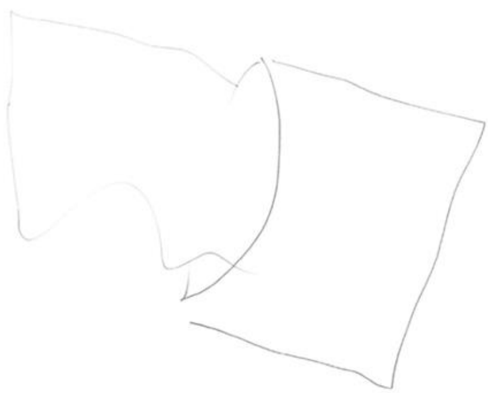


Users place food waste in bin. It is then hand ground and users can spread on vegitation before leaving

Personal refuse sack

Sc_C19

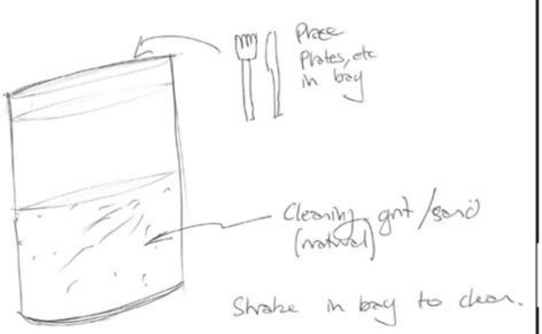
Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



Tent elevated off ground - less damage to plants on ground	Vacuum bag suction bin bag Put your rubbish in this bag, create a vaccuum, easy to carry rubbish around with you - light and small

Sc_C20

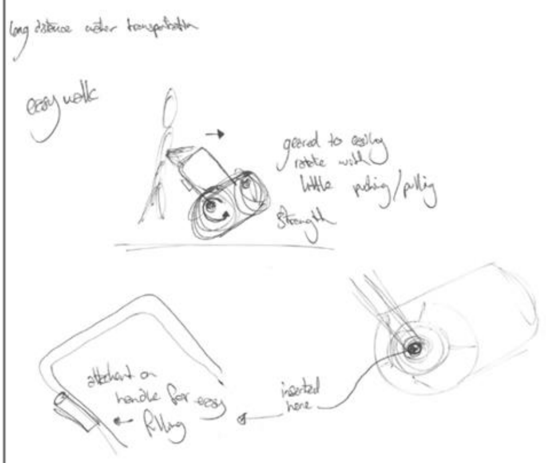
Original brief: Camping is a popular activity but can have negative environmental impacts through disruption to wildlife; litter and pollution of water sources. Generate concepts for products that reduce the negative impacts of camping.



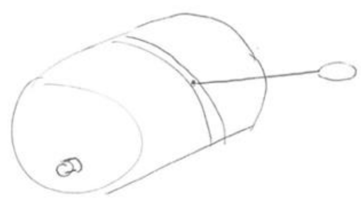
Shake to clean - no soaps!	the mic records all the noise you and your mates make @ campsite and then the speakers produce the mirrored wavelength to make it nice & quiet for the wildlife around

Sc_C21

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



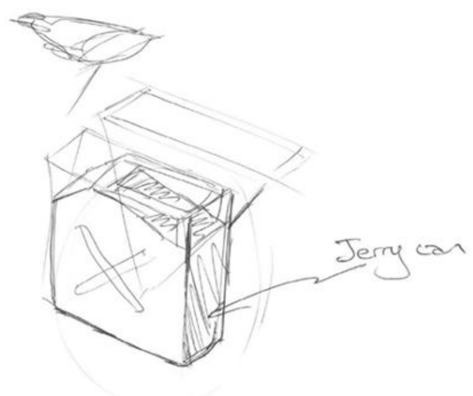
Water barrel with attached handle for pushing water over long distances. For those who need to walk to collect water



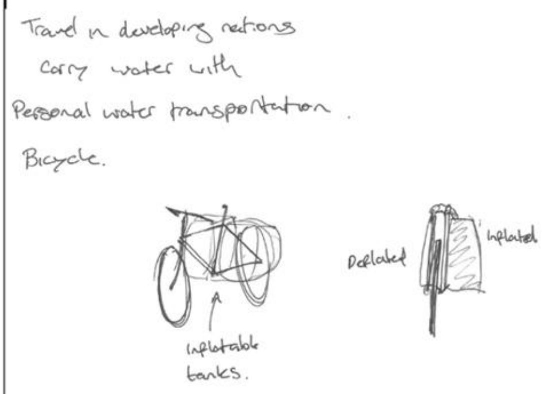
Barrell with string attached to ball-bearing ring that can be easily dragged / rolled when full

Sc_C22

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Jerry can panier for bicycles

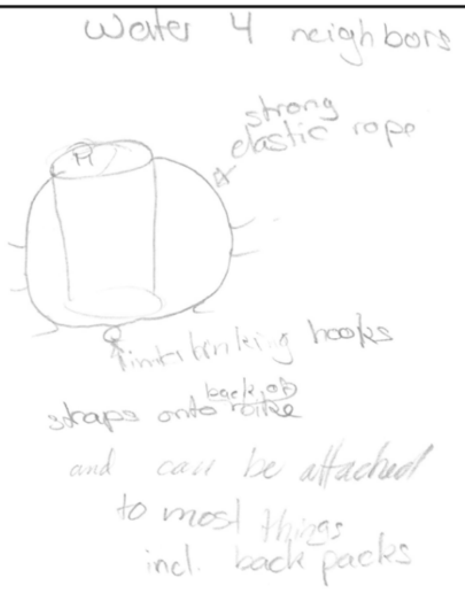


Tank either side of rear wheel.
Hooks on to conventional bicycle

Sc_C23

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

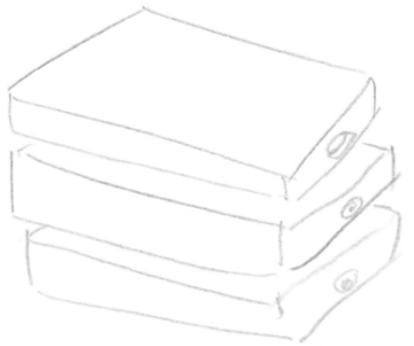
Water container bag, to make it easier to carry water



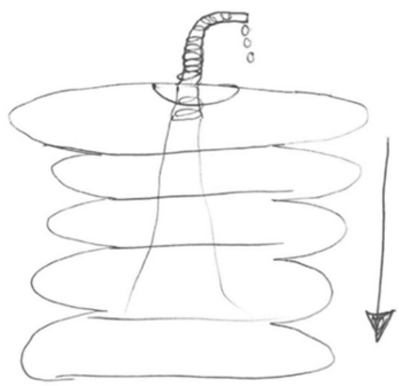
Can be transported / used by anyone

Sc_C24

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Stackable square water tanks which are relatively light weight and easy to transport



Large foldable bottles could be sent that people would fill up and the water would then be filtered cleanly through a pump

Sc_C25

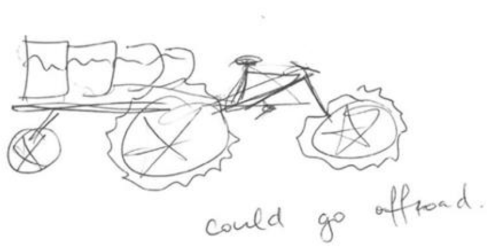
Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

Bus-mounted water transport

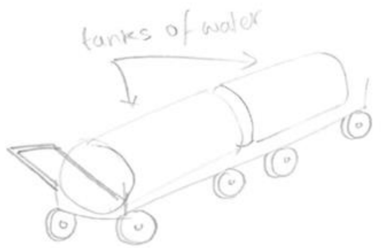
Heavy lift drones for mass delivery

Sc_C26

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



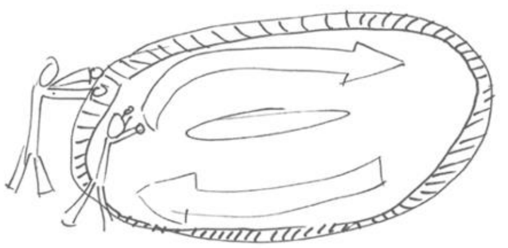
cargo bike

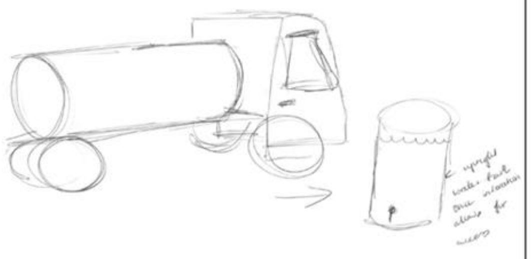


Using a pull bar and a lot of tough thick tyre treads for long rough terrain it could be easy for 1/2 people to transport a large amount of water

Sc_C27

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

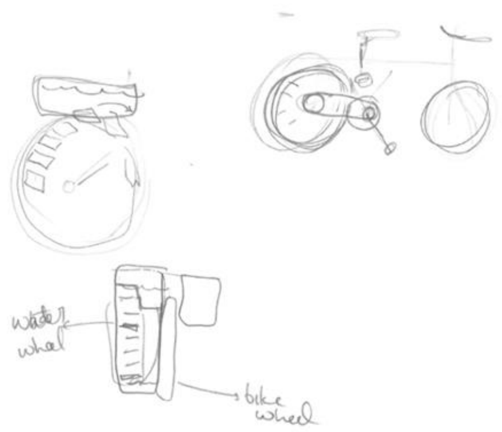


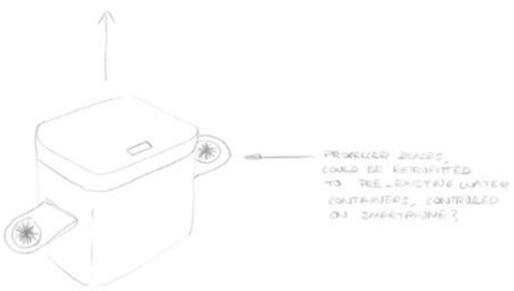


A large water balloon that is a hoop so when pushed, it travels like a tank track. Thin tank tracks are on the outside to protect innards from rough terrain	Large water storage tank that can be moved by lorry to places that need it Upright water tank once in location allows for access

Sc_C28

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

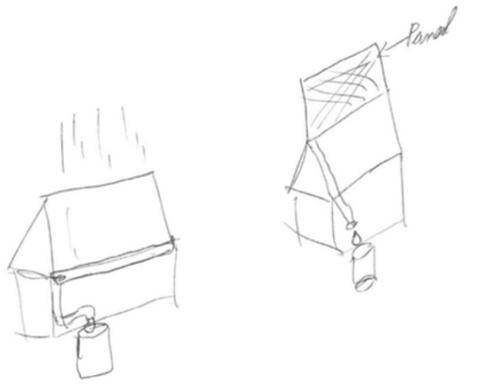
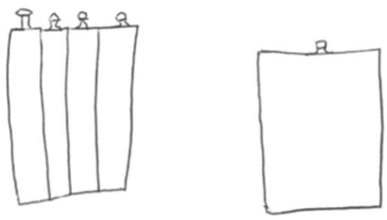




Bike that uses H20 for conservation of momentum (like a water wheel) - once bike is in motion it can move a long ways w/ less effort	Solar powered water drones that are cheap and easy to look after, clean, repair, robust stucture

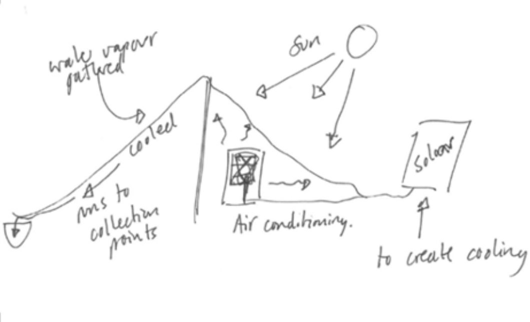
Sc_C29

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

	
Collection of rain water Panel that retains moisture to collect it	Individual water bottles are all rectangular shaped to be easily transported in a box, lots can be stored together

Sc_C30

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.

	
Designed to be filled and used over varying conditions. Obviously consideration must be given to the max. weight people can actually push, relative to the capacity of the holder	No transportation Made at source from water in atmosphere

APPENDIX 5 – CODING FOR SIMILARITY EXPERIMENTS

APPENDIX 5A – EXAMPLE INFORMATION SHEET

An example of an information sheet given to participants. Minor modifications were made at different phases of the research to reflect different participant demographics.

University of Strathclyde Glasgow

Participant information sheet for similarity and concept creation study

Name of department: Design Manufacture and Engineering Management
Title of the study: Cognitive processing in conceptual design: similarity and concept creation

Introduction
This research is being conducted as part of research towards fulfilment of a PhD. The researcher is Christopher McTeague, a PhD candidate within the Design Manufacture and Engineering management (DMEM) department at Strathclyde University.

What is the purpose of this investigation?
The purpose of this investigation is to investigate the cognitive processes involved in two different tasks: similarity judgements and a concept creation task.

Do you have to take part?
The experiment is entirely voluntary. You are not required to participate in the experiment and you are free to withdraw from the experiment at any time, both before and during the data recording session.

What will you do in the project?
You will be asked to do one of the following:
(i) make judgements about pairs of design concept sketches and enter data into a computer, or
(ii) create design concepts in response to a brief by following task instructions
Following the experiment you will be asked to answer some questions about the study verbally. These answers will be recorded.

Why have you been invited to take part?
You have been invited to take part in this study because you have at least 1 years of formal Product Design education

What are the potential risks to you in taking part?
The researcher and departmental ethics committee have determined that there are no risks involved in participating in this study.

What happens to the information in the project?
You will be asked to provide your name and some basic demographic information (gender, department). In addition, a comprehensive record of the work conducted during the session will be kept. For example the answers to questions or concepts created. The information gathered from the experiment will remain entirely confidential and will be stored in accordance with Strathclyde University's data management requirements.

The University of Strathclyde is registered with the Information Commissioner's Office who implements the Data Protection Act 1998. All personal data on participants will be processed in accordance with the provisions of the Data Protection Act 1998.

Thank you for reading this information – please ask any questions if you are unsure about what is written here.

What happens next?
If you are happy to be involved in the experiment, you will be asked to sign a consent form confirming this.

If you do not wish to be involved in the experiment then nothing further is required, thank you for your time and attention.

Researcher contact details:
Christopher McTeague
PhD Student
christopher.mcteague.@strath.ac.uk

DMEM Department
University of Strathclyde
Architecture Building
131 Rottenrow
Glasgow
G4 0NG

If you have any questions/concerns, during or after the investigation, or wish to contact an independent person to whom any questions may be directed or further information may be sought from, please contact:

Secretary to the University Ethics Committee
Research & Knowledge Exchange Services
University of Strathclyde
Graham Hills Building
50 George Street
Glasgow
G1 1QE

Telephone: 0141 548 3707
Email: ethics@strath.ac.uk

Consent Form for Similarity and Concept Creation study

Name of department: Design Manufacture and Engineering Management
Title of the study: Design concept similarity experiment

- ☐ I confirm that I have read and understood the information sheet for the above project and the researcher has answered any queries to my satisfaction.
- ☐ I understand that my participation is voluntary and that I am free to withdraw from the project at any time, without having to give a reason and without any consequences.
- ☐ I understand that I can withdraw my data from the study at any time.
- ☐ I understand that any information recorded in the investigation will remain confidential and no information that identifies me will be made publicly available.
- ☐ I consent to being a participant in the project
- ☐ I consent to having my voice recorded in the post-experiment interview

(PRINT NAME)	
Signature of Participant:	Date:

APPENDIX 5B – EXAMPLE CONSENT FORM

The demographic form asked for details relevant to three independent experiments conducted by the research and two colleagues

Design behaviour experiments

Participant demographic sheet

Note: all data is held in strict confidentiality

* = only enter if applicable

Personal details	
Forename	
Surname	
Gender (M/F/Other)	
Date of Birth	
Occupation (Student / Specify other)	
Student number	
Email address	
Current student details (all students)	
Department	
Degree	
Year of study	
Previous higher education (For current postgraduate students only)	
Previous department	
Previous degree	
Highest degree award attained	
Duration of study	
Additional notes	

Please turn over

Design/CAD experience	
If you used CAD software in the past which one was it?	
Years of experience in the CAD software listed above	
How many years of design experience do you have (outside of university)	
Are you left handed, right handed or ambidextrous?	

APPENDIX 5C – CODING HANDBOOK FOR COMMONALITY AND DIFFERENCE CODING ROUND 1

1. Overview

This document contains instructions for coding participant responses to a 'commonality and difference listing' task. Experiment participants were asked to identify commonalities and differences between pairs of design concepts. The responses must be coded so they can be counted, and since there are two types of differences, multiple ratings are needed to ensure consistency.
Commonality and difference listing: Participants were given 30 pairs of design concepts and an instruction sheet which asked them to list *either* the commonalities or differences for each pair. As part of the experimental analysis, it is necessary to count the number of commonalities and differences listed by each participant. Counting commonalities is straight forward but there are two kinds of differences and some interpretation is needed to code these. You have received this document because you have kindly agreed to participate in this exercise as an independent coder.
It is important that you remain naïve to the aim of the experiment, but all the necessary information to complete the coding will be provided in the following section.
e: christopher.mcteague@strath.ac.uk
t: +44(0)141 574 5293

2. Concept coding and ranking process

2.1. Coding procedure

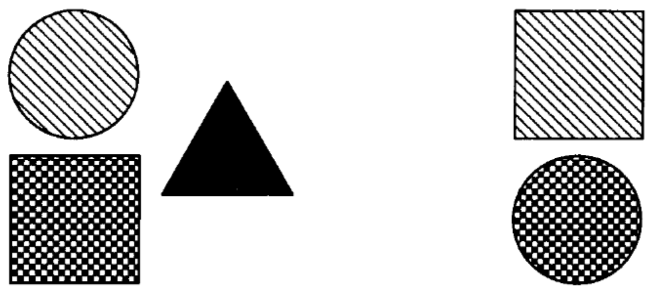
You have been given an excel spreadsheet with the data to be coded. Each row on the spreadsheet represents the differences listed by one participant in response to one pair of design concepts. You have also been given a copy of all the pairs used in the experiment. The response rows are numbered 1-30 so that you can view the responses alongside the stimuli given to the participants. The data given to you is a partial sample of the full response set so it is possible that not all concepts will appear.
You are asked to read the instructions below which define the two types of differences, and then code each cell as one of these types by entering a code in the cell below.

2.2. Differences

There are two types of differences

- **Alignable differences** are those differences that are related to commonalities between two concepts
- **Nonalignable differences** are those differences that are unrelated to the commonalities between two concepts

Example: In the figure below, the two drawings share the common relation 'x *above* y'. In other words, both drawings have one shape which is *above* another shape, this is a commonality.



(left) (right)

The fact that there is a circle on top (left) vs a square on top (right) is an **alignable difference**. A participant may phrase this of the form: *"there is a circle on top in one and a square on top in the other".* The difference is **alignable** because the difference (circle vs square) is related to the commonality (the *above* relation).
The triangle in the left drawing is a **nonalignable difference.** It is not related to any corresponding commonality on the right. This could be phrased *"there is a triangle in one configuration but not in the other*".
Another way to view it is: alignable differences are corresponding elements which are unlike, whereas nonalignable differences are non-corresponding elements.

2.3. Scoring instructions

You are tasked with coding each cell as either an alignable difference (AD) or a nonalignable difference (ND), superficial (S), or other (O)

<u>Identifying differences</u>

Alignable differences - Participants need to make explicit or implicit mention of a difference value along some common dimension for both concepts. All of these examples are alignable differences:

- "A police car is a car, an ambulance is a van"
  - *An explicit mention. Both are vehicles, but are different types*
- "Police cars and ambulances are different kinds of vehicles"
  - *An implicit mention. Booth are vehicles, but are different kinds*
- "A hotel is expensive; a motel is cheap"
  - *The participants mention contrasting properties of the two items*
- "A hotel is more expensive than a motel"
  - *The participant used an explicit comparative construction*
- "One concept has wheels, the other has ski's"
  - *The concepts differ along some common dimension. They both have parts which interact with the ground to facilitate transport, but they have different kinds.*

Nonalignable differences - All other differences not covered by the above are nonalignable differences,

- "A police car has weapons in it, an ambulance does not"
  - *No common counterpart is mentioned in the case of the ambulance*

- "A watch has a face, but a necklace does not"
    - *There is no common dimension mentioned in the case of the necklace*
- "You read a magazine, but you do not read a kitten"
    - *Negation of one items property as applied to the other*
- "Concept A has a solar panel"
    - *Mention of one concept, but not the other*

This is a conservative measure of alignable differences because participants may say "Cars have four wheels, and motorcycles don't" when what they really mean is "Cars have four wheels, motorcycles have two". In this case the response is coded as nonalignable, even if the participant was thinking about an alignable difference. This issue has been anticipated and should not influence your coding. **It is important that your interpretation is based purely on the content of the response. This is especially important if you look at the accompanying stimuli-pair. You must code the nature of the participants response, rather than any notion of the "true" alignability of the response.**

Superficial differences: the participants have been tasked with listing differences related to the concept, but not the sketch itself. They were asked to focus on what the sketch represented, rather than making comments of the sort:

- "one is a tidy sketch, but the other is rough".

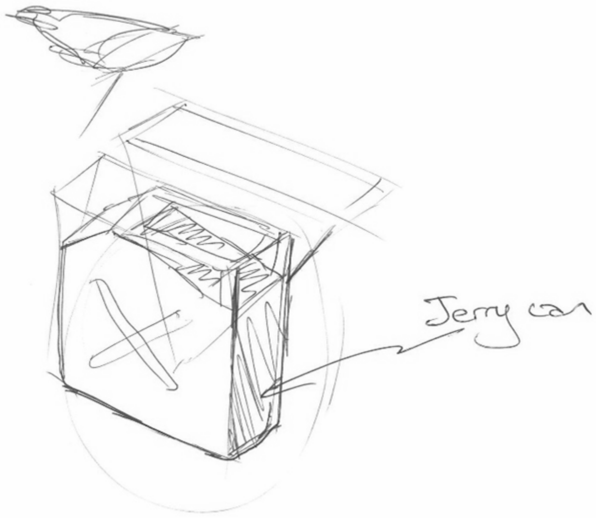
If any such comments are identified they should be marked as superficial (S).

Other (O): if any response cannot be marked as AD, ND or S, please mark it as 'O' and if possible explain the issue.

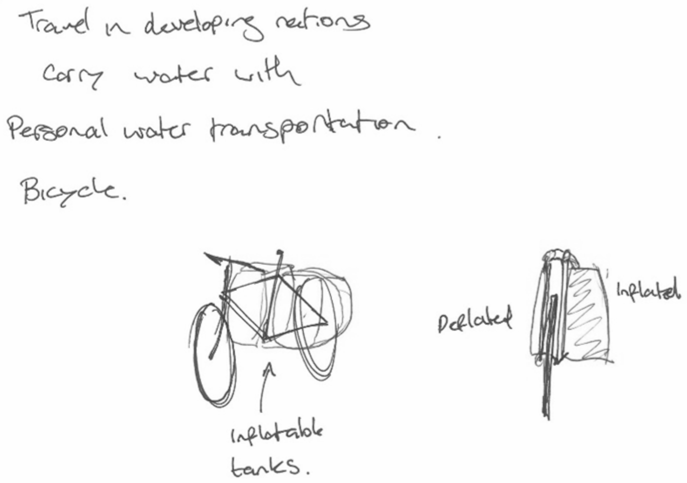
3. Stimuli

An example of the type of stimuli given to the participants in the listing task is shown on the next page. The original brief is presented alongside the pair of concepts. In the original experiment the participants were asked to provide annotations in a text box; these have been re-typed for clarity below each sketch.

Original brief: Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.



Jerry can panier for bicycles



Tank either side of rear wheel.
Hooks on to conventional bicycle

4. Summary sheet

Your task: code each cell as alignable difference (AD), nonalignable difference (ND), superficial response (S), or other (O).

- **Alignable difference (AD)** – the difference is related to a commonality
- **Nonalignable difference (ND)** – the difference is unrelated to any commonality
- **Superficial response (S)** – the participant has commented on some aspect of the sketch, or some other superficial aspect which does not relate to the idea represented by the stimuli
- **Other (O)** – the response cannot be coded with any of the above codes.

APPENDIX 5D – CODING HANDBOOK FOR COMMONALITY AND DIFFERENCE CODING ROUND 2

1. Overview

This document contains instructions for coding participant responses to a 'commonality and difference listing' task. Participants were given 30 pairs of design concepts and an instruction sheet which asked them to list *either* the commonalities or differences for each pair. As part of the experimental analysis, it is necessary to count the number of commonalities and differences listed by each participant. Counting commonalities is straight forward but there are two kinds of differences and some interpretation is needed to code these. You have received this document because you have kindly agreed to participate in this exercise as an independent coder.

It is important that you remain naïve to the aim of the experiment, but all the necessary information to complete the coding will be provided in the following section.

e: christopher.mcteague@strath.ac.uk
t: +44(0)141 574 5293

2. Concept coding process

2.1. Materials

You have been given an excel spreadsheet with the data to be coded. Each row on the spreadsheet represents the differences listed by one participant in response to one pair of design concepts. Participants were instructed to list each difference on a new line, and these differences have been transformed into rows for easier processing. Each line in the initial response corresponds to one cell in excel.

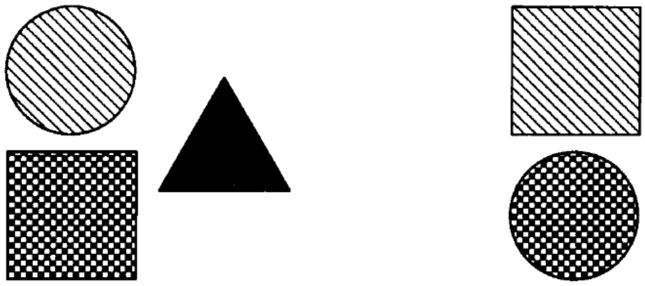
You have also been given a copy of all the pairs used in the experiment. The response rows are numbered 1-30 so that you can view the responses alongside the stimuli given to the participants. The data given to you is a partial sample of the full response set so it is possible that not all concepts will appear.

2.2. Differences

There are two types of differences

- **Alignable differences** are those differences that are related to commonalities between two concepts
- **Nonalignable differences** are those differences that are unrelated to the commonalities between two concepts

Example: In the figure below, the two drawings share the common relation 'x *above* y'. In other words, both drawings have one shape which is *above* another shape, this is a commonality.



(left) (right)

The fact that there is a circle on top (left) vs a square on top (right) is an **alignable difference**. A participant may phrase this of the form: “*there is a circle on top in one and a square on top in the other”.* The difference is **alignable** because the difference (circle vs square) is related to the commonality (the *above* relation).

The triangle in the left drawing is a **nonalignable difference.** It is not related to any corresponding commonality on the right. This could be phrased “*there is a triangle in one configuration but not in the other*”.

Another way to view it is: alignable differences are corresponding elements which are unlike, whereas nonalignable differences are non-corresponding elements.

2.3. Scoring instructions

You are tasked with coding each cell with one of five codes. The first two codes are the two codes hat describe the type of difference that the participant has listed:

- **Alignable differences (AD) –** These differences are typically stated as explicit or implicit mention of a difference value along some common dimension for both concepts.
- **Nonalignable differences (ND) –** These differences are not related to commonalities. They can be identified when a fact is stated about one item but negated or ignored for the other.

These additional codes are used to capture responses that are superficial, formulaic or which for some reason cannot be coded as an alignable or nonalignable difference:

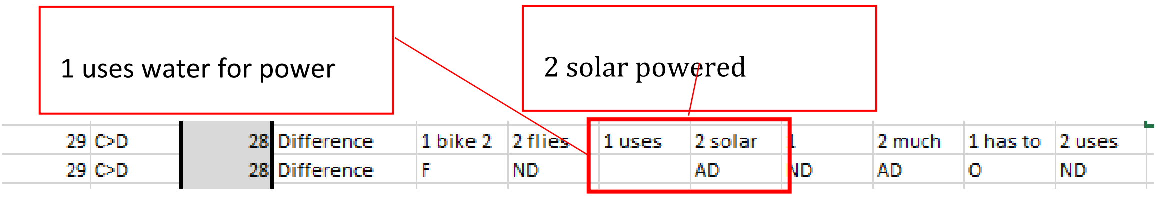
- **Superficial / sketch differences (S) -** the participants have been tasked with listing differences related to the concept, but not the sketch itself. If the response refers to some aspect of the sketch it should be coded as ‘superficial’, e.g:
    - *“one is a tidy sketch, but the other is rough”.*
- **Formulaic descriptions (F) -** if the participant has simply described both concepts without pulling out a meaningful difference, code it as ‘F’ for formulaic:
- **Other (O):** if any response cannot be marked as AD, ND, S, F, or if you are unsure about how it should be coded, please mark it as ‘O’ and if possible explain the issue.

2.4. Coding procedure

Based on the previous definitions and additional examples presented in this book, please code each of the responses in the spreadsheet by placing a code in the cell directly beneath that code.

1) It is important that your interpretation is based purely on the content of the response. This is especially important if you look at the accompanying stimuli-pair. You must code the nature of the participant's response, rather than any notion of the "true" alignability of the response.
2) If you have trouble deciding between alignable and nonalignable differences, ask yourself: is there a common dimension which relates the two facts? If you cannot find a plausible common dimension, then code ND.
3) Finally, the participants were asked to list each difference on a new line. In most cases alignable differences are placed on one line, e.g. "the left concept carries water, but the right concept carries people", however in some cases alignable differences are split over two cells.
   a) If you infer that sequential cells are an alignable difference that has been split over two lines, code AD in the second of the two related concepts and leave the first blank (see below).
   b) If you infer that the participant has listed alignable differences *that are more than one cell apart,* **do not code AD**

Example, in the image below the two highlighted statements are inferred to be related and are coded as 'AD' in the right-most cell.



Please use a conservative measure of alignable differences because participants may say "Cars have four wheels, and motorcycles don't" when what they really mean is "Cars have four wheels, motorcycles have two". In this case the response is coded as nonalignable, even if the participant was thinking about an alignable difference. This issue has been anticipated and should not influence your coding.

2.5. Examples from psychology literature

Alignable differences – Alignable differences can be phrased in many ways. They are frequently stated as an explicit or implicit mention of a difference value along some common dimension for both concepts. Pay particular attention to the second white bullet below each example. The format of these sentences is a useful way to help you interpret the response.

- *"A police car is a car, an ambulance is a van"*
    - An explicit mention of an alignable difference.
    - Both are vehicles (common dimension) but are different kinds (difference related to common dimension).
- *"Police cars and ambulances are different kinds of vehicles"*

  - An implicit mention of an alignable difference.
  - Both are vehicles (common dimension) but are different kinds (difference related to common dimension).
- *"A hotel is expensive; a motel is cheap"*
  - The participants mention contrasting properties of the two items.
  - Both have relative appraisals of value (common dimensions) but are different.
- *"A hotel is more expensive than a motel"*
  - A common dimension (cost) is contrasted along a varying magnitude.
  - As in the previous example, both have a cost (common dimension) but one is greater than the other.
- *"One concept has wheels, the other has ski's"*
  - The concepts differ along some common dimension.
  - They both have parts which interact with the ground to facilitate transport, but they have different kinds.
- *"Tigers hunt mammals and sharks hunt fish"*
  - In this case there is a common relation (hunt) between different pairs of entities.
  - In both pairs a hunter is hunting a pray (common dimension) but the examples of the hunter and the pray are different.

Nonalignable differences – These differences are not related to commonalities. They can be identified when a fact is stated about one item but negated or ignored for the other.

- *"A police car has weapons in it, an ambulance does not"*
  - A fact is asserted about one item, but denied or negated for the other.
- *"A watch has a face, but a necklace does not"*
  - Again, an explicit negation of something indicates a nonalignable difference.
- *"You read a magazine, but you do not read a kitten"*
  - Negation of one item's property as applied to the other
- *"Concept A has a solar panel"*
  - Mention of one concept, but not the other. This should be treated the same as a negation.
- *"Tigers hunt mammals and dogs chase cats"*
  - In this example there are no commonalities, the entities are different (the animals) as are the relations (hunt and chase)

2.6. Notes to help you interpret responses in design

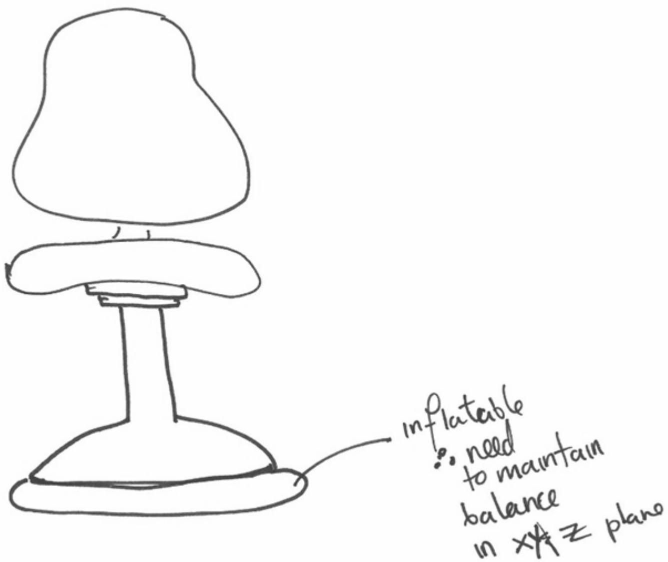
- If the participant provides and explanation for a difference or an implication of a difference, count only the difference and ignore the cause / implication
  - *"_ The first concept has a bag [difference] which will keep the product clean and improve the user experience [implication]."*
- Some differences contrast variables with binary values, these may seem like negations but are actually describing alignable differences
  - *"'One involves a expandable/collapsible storage, the other is a fixed volume"*
- Alignable differences can manifest as differences in the number of entities. Generally, if the response includes a relative description such as: more, greater, cheaper, faster… the response is an alignable difference.

    - *“one may require more than one person”*
- Some responses which include speculation, simulation or assumptions can be coded
    - *“Right might be harder to get water out of.”*
- Some statements appear to contrast two differences, but it is not possible to find a plausible common relation between them. These may be nonalignable (ND) or formulaic (F) depending on your interpretation.
    - *“One requires mechanical effort to actuate, the other is a storage system”*
- If a participant lists two differences within the one statement:
    - If both differences are AD, code AD
    - If both differences are ND, code ND
    - If the statement contains an AD and ND, code as ‘other (O)

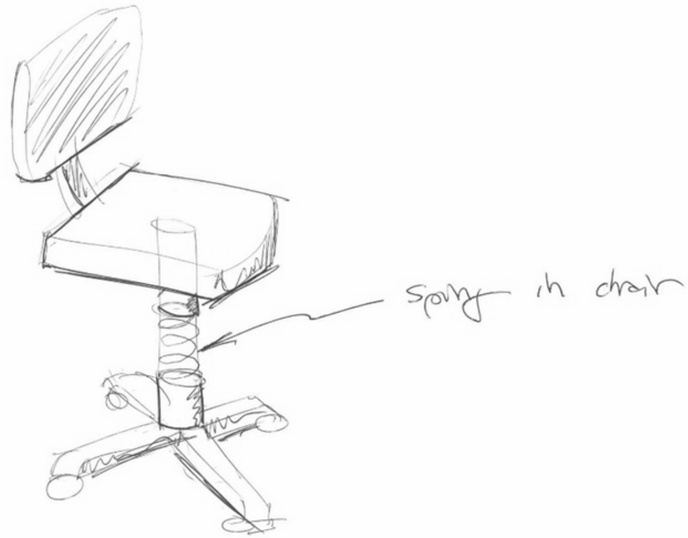
3. Stimuli

An example of the type of stimuli given to the participants in the listing task is shown on the next page. The original brief is presented alongside the pair of concepts. In the original experiment the participants were asked to provide annotations in a text box; these have been re-typed for clarity below each sketch.

Original brief: Sitting in the same position for long periods may be harmful to health. Generate concepts for products that may facilitate physical exercise whilst completing activities in a seated position in the home and office.



Seat that moves and you need to adjust your position / maintain balance



Bounce your way to fitness

4. Summary sheet

Your task: code each cell as alignable difference (AD), nonalignable difference (ND), superficial response (S), or other (O).

- **Alignable difference (AD)** – the difference is related to a commonality
- **Nonalignable difference (ND)** – the difference is unrelated to any commonality
- **Superficial response (S)** – the participant has commented on some aspect of the sketch, or some other superficial aspect which does not relate to the idea represented by the stimuli
- **Formulaic description (F)** – the participant has simply described the two concepts
- **Other (O)** – the response cannot be coded with any of the above codes.

Remember:

- Base your interpretation on the content of the response
- Ask yourself: is there a plausible common dimension?
- Use a conservative measure (favour ND over AD)
- Sequential statements with plausible common dimensions can be listed as AD

Appendix 6 – Coding for combination experiments

Appendix 6A – Combination coding scheme v0

Table A6-1 - Combination coding scheme v0

Success	**Response type**	**Codes**	**Code descriptions**	**Collapsed codes**
Design concept successfully created	Combination	F – Feature mapping	A single artefact comprising a mixture of features from both base concepts.	F
		R – Relational	The two base concepts have been related together by an external relation. The relation can be structural, behavioural, functional, or some other abstract external relation. May appear as two, physically distinct artefacts or (in the case of a structural relation) a single artefact in which the two base concepts are clearly identifiable beyond the minor modification required.	R
		E – Entity substitution	A design artefact or other artificial entity in one concept has been substituted for an entity in the other design concept.	
		A - Ambiguous	A design concept that can be attributed to an act of combination, but is not unambiguously a featural or relational combination.	A
	Non-combination	ML – Modification of left concept	A new design artefact that shares commonalities with the left base concept, but none of the features can be attributed to the right concept.	NC
		MR – Modification of right concept	A new design artefact that shares commonalities with the right base concept, but none of the features can be attributed to the left concept.	
		MCB – Modification based on commonalities of both concepts	A new design artefact that shares elements of both base concepts, but has unique elements of neither. As if the designer has started with elements common to both artefacts as a starting point for generating something new.	
		G – Generated new concept	A new design concept has been created that shares no identifiable commonalities with either of the two base concepts.	
Design concept not successfully created	Other	I – Insufficient information to interpret sketch	The representation of the design concept cannot be interpreted.	I
		U – Unclear how concept addresses brief	A new design concept has been created. It is possible to infer the intent of the designer but the concept does not appear to address the design brief.	U
		NR – No response (could not generate)	A participant attempted the trial but did not produce a new design concept. This does not include trials that were not completed due to finishing the experiment early, i.e., trials at the end of the participants session.	---

APPENDIX 6B – COMBINATION CODING SCHEME V1

Table A6-2 -

Code	Label	Code description
Successful trial - Combination		
Featural	F	In featural combinations, the newly created design artefact contains elements of both of the artefacts from both stimulus concepts that have been transferred or mixed together to create a new concept. 'Element' refers to any aspect of the design artefact but *not* the representation of the artefact such as the sketch quality, line types or shading.
Relational	R	Relational combinations are new artefacts or systems of artefacts that contain the design artefacts from both stimulus concepts joined together by some external relation. The external relation may be physical or abstract (e.g., spatial, functional). The original design concepts may have been modified to facilitate the creation of a new relation between them, but you should be able to identify that the essences of both original concepts are still present.
Ambiguous	A	Ambiguous combinations are those that can be coded as both featural or relational.
Successful trial - Non-combinational ideation		
Non-combination	NC	A non-combinational design concept contains a design artefact that can be understood and addresses the brief but wherein it is not clear that combination occurred. Examples may appear to be entirely new concepts, modifications of one base concept (without elements of the other), or a new concept that builds on commonalities of both base concepts (but without any transfer of elements that would indicate a featural combination).
Unsuccessful trial		
Other	I	Insufficient info - The representation of the design concept cannot be interpreted
	U	Unclear how address brief - A new design concept has been created. It is possible to infer the intent of the designer but the concept does not appear to address the design brief.
	O	Other – the concept cannot be coded by any of the other codes

APPENDIX 6C – COMBINATION CODING HANDBOOK

1. Overview

This document contains instructions for coding participant responses to a 'design concept combination' task. Under experimental conditions, product design engineering students were given pairs of design concepts that had been created previously in response to a design brief. They were tasked with combining them to create new design concept that addresses the same brief. To analyse the results of the experiment, it is necessary to count the different types of combined concepts that were produced, and independent judges are need to help in the process of coding the concepts into types.

You have received this document because you have kindly agreed to participate in this exercise as an independent coder. You will remain naïve to the aim of the experiment, but the document contains all the necessary definitions, instructions and examples for you to carry out the coding process.

e: christopher.mcteague@strath.ac.uk
t: +(0)7870358776

2. The concept combination task

This section describes the experiment that the participants took part in and provides definitions for the stimuli and outputs.

Experiment. Product Design Engineering students were presented with 30 stimuli. Each stimulus contained a pair of design concepts and the design brief that they have been created in response to. Participants were presented with the stimuli, one at a time, and were asked to combine the design concepts to create a new design concepts that addresses the same brief.

2.1. Stimuli: pairs of design concepts

An example of the stimulus presented to the experiment participants is shown in Figure 5-1. The following terms are used to define the stimuli and its elements. Stimuli contain pairs of **design concepts**, and **design concepts** contain **design artefacts** as well as **generic entities and existing products**.

Stimulus – a pair of design concepts, presented side by side with a design brief written above it. This is shown in Figure 5-1.

- **Design concept** – a proposal for a product that addresses the requirements specified in the design brief. In this document, all design concepts are represented by sketches, text-based descriptions and annotations. All design concepts include a design artefact (defined below) and may also include other things termed 'generic entities and existing products' (defined below).
    - **Design brief** – A statement that describes a scenario that has some problem associated with it, and the instruction to generate concepts that address the problem stated in the brief.

- **Design artefact(s)** – this is the thing(s) being designed. 'Artefact' means 'artificial object'. It is the designers' proposal for a new product that will address the requirements contained in a design brief.
- **Generic entities and existing products** – 'generic entity' is used to refer to anything represented in the design concept that is not the design artefact. Some design concepts contain additional entities such as existing products, humans (users), environments etc. Existing products are those which one would expect a user to own or have access to already, such as a bicycle, smartphone, car or backpack.

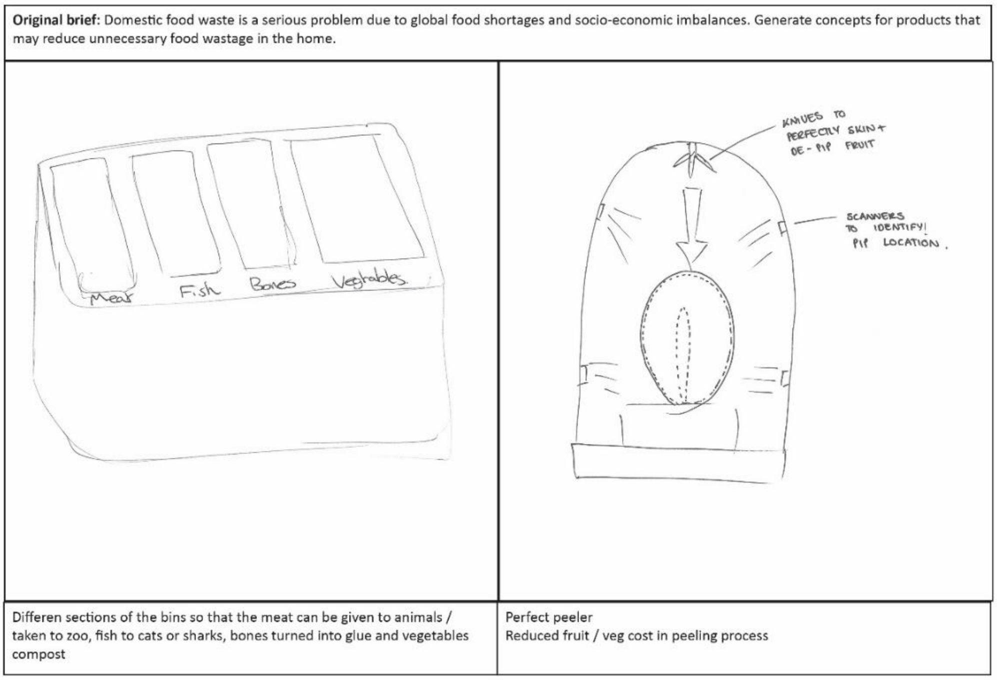


Figure 1 - Example of a single stimulus containing a pair of design concepts, i.e., the 'base concepts' for the experiment.

Example 1: A stimulus is shown in Figure 5-1.

The design brief states: "Domestic food waste is a serious problem due to food shortages and socio-economic imbalances. Generate concepts for products that may reduce unnecessary food wastage in the home."

Consider the design concept on the right-side of Figure 5-1. This design concept represents something described as a 'perfect peeler'. It may be inferred that if such a product were to exist, people could peel their fruit with minimal waste, thereby reducing the unnecessary food waste in the kitchen, and thus satisfying the brief. In this case, the design artefact is the product that does the peeling. That is, a product with a base and surrounding dome, including scanners and knives. The circular form in the middle of the artefact may represent some kind of fruit, which would be a generic entity.

Example 2: A different stimuli is shown in Figure 2

The design brief states: “Long distance water transportation may be necessary in drought-prone developing nations but can be problematic due to lack of resources and infrastructure. Generate concepts for products that may facilitate water transportation in developing nations.”

Consider the design concept on the left-side of Figure . This design concept represents something described as a ‘jerry can panier for bicycles’. It may be inferred that if such a product were to exist, people could fill a jerry can with water, store it on their bicycle, and thus the product would facilitate water transportation in developing nations. In this case, it may be inferred that the design artefact is a panier (i.e., a carry mount or frame that attaches to the bicycle and holds the jerry can), the bicycle and jerry can would be existing products, and the water would be a generic entity.

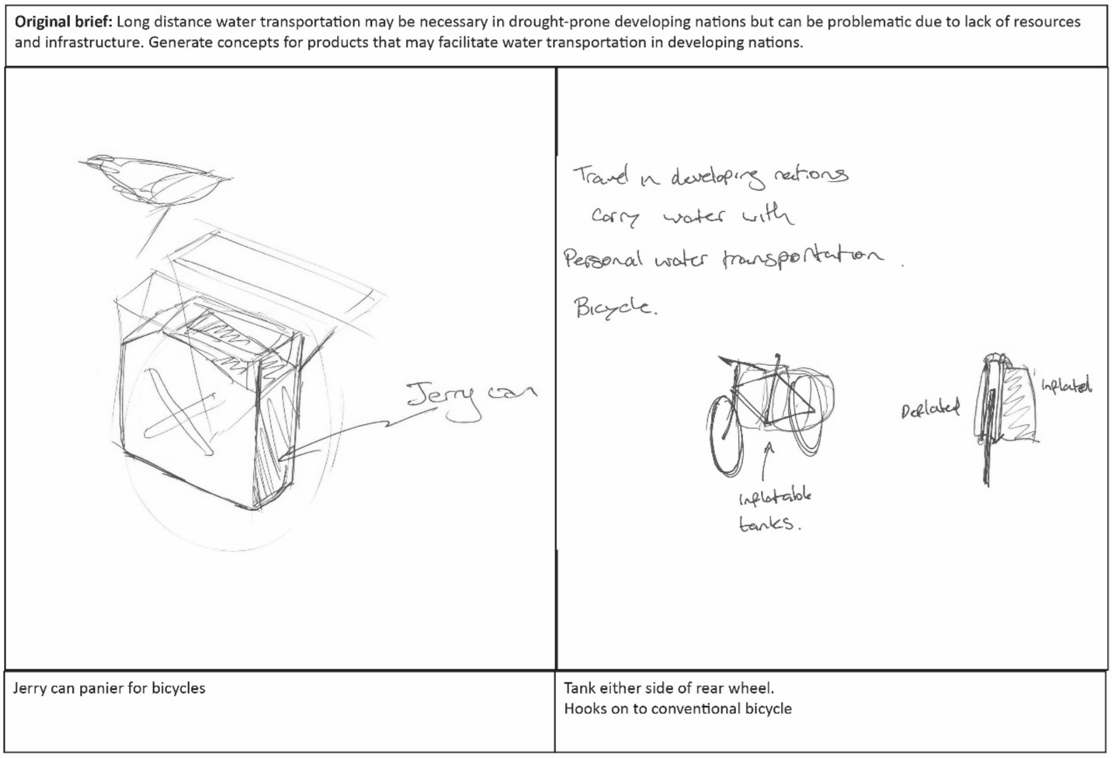


Figure 2 – A further example of a stimulus to illustrate how design artefacts can be shown alongside existing products.

2.2. Output: combined concepts

Figure shows a new design concept that has been created by combining the two design concepts contained in the stimulus in Figure 5-1. The design concepts created by the participants adhere to the same definition of ‘design concept’ provided in the previous section. It may be inferred that the designer has combined the two design concepts by modifying the bin (Figure 5-1 left) to act as a receptacle for the waste produced by the peeler (Figure 5-1, right). Instructions for how to code this kind of combination and supporting examples are provided in the subsequent sections.

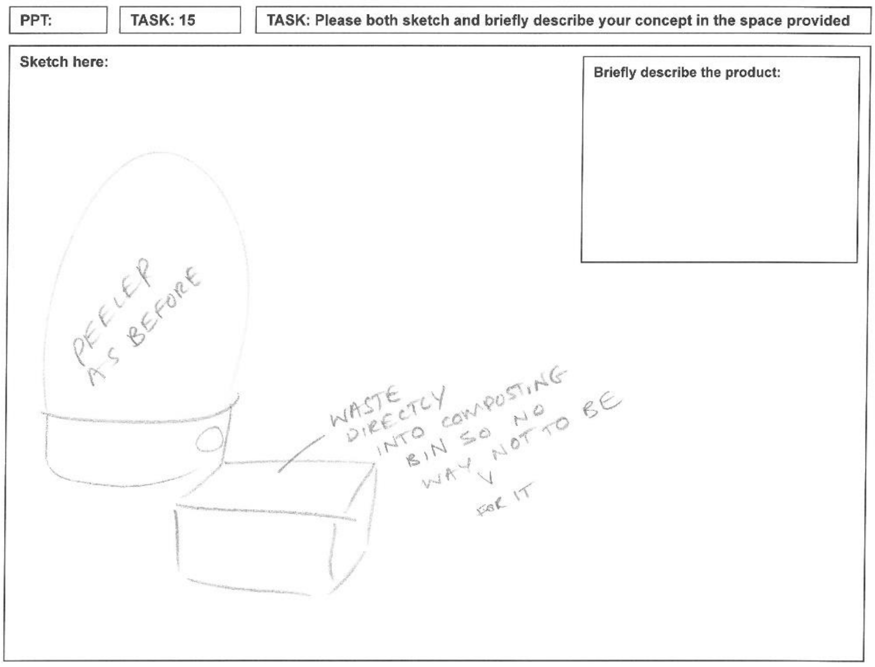


Figure 3 – An example of a combined design concept, created in response to the design brief and design concepts shown in Figure .

3. Concept coding process

This section describes the materials you have been given, the codes that are to be applied to the concepts, and the procedure for applying them.

3.3. Materials

You have been given the following materials:

- The pairs of design concepts presented to the participants in their combination task, these are the 'base concepts'.
- The combined sketches produced by the designers
- An excel spreadsheet for entering your codes. Each row of the spreadsheet indicates the combined concept and corresponding pair of base concepts.

The data given to you is a partial sample of the full response set so it is possible that not all input concepts will appear.

3.4. Combination codes

You will code the outputs of the combination task by using one of seven codes. Three of the codes describe new design concepts created through combination, one describes new design concepts created through some form of ideation other than combination and three capture unsuccessful responses. You will examine each concept, compare it with the base concepts and determine which code is the most applicable. The codes are shown in Table and the coding procedure is shown in Section 3.5.

Table 1 – The list of codes for coding combination types

Code	Label	Code description
Successful trial - Combination		
Featural	F	In featural combinations, the newly created design artefact contains elements of both of the artefacts from both stimulus concepts that have been transferred or mixed together to create a new concept. ‘Element’ refers to any aspect of the design artefact but *not* the representation of the artefact such as the sketch quality, line types or shading.
Relational	R	Relational combinations are new artefacts or systems of artefacts that contain the design artefacts from both stimulus concepts joined together by some external relation. The external relation may be physical or abstract (e.g., spatial, functional). The original design concepts may have been modified to facilitate the creation of a new relation between them, but you should be able to identify that the essences of both original concepts are still present.
Ambiguous	A	Ambiguous combinations are those that can be coded as both featural or relational.
Successful trial - Non-combinational ideation		
Non-combination	NC	A non-combinational design concept contains a design artefact that can be understood and addresses the brief but wherein it is not clear that combination occurred. Examples may appear to be entirely new concepts, modifications of one base concept (without elements of the other), or a new concept that builds on commonalities of both base concepts (but without any transfer of elements that would indicate a featural combination).
Unsuccessful trial		
Other	I	Insufficient info - The representation of the design concept cannot be interpreted
	U	Unclear how address brief - A new design concept has been created. It is possible to infer the intent of the designer but the concept does not appear to address the design brief.
	O	Other – the concept cannot be coded by any of the other codes

Tips for selecting codes. A key distinction between featural (F) and relational (R) codes relies on the identification of the ‘boundary’ that distinguishes the design artefact(s) from the other entities.

- Featural combinations involves taking an element of the artefact (i.e., something 'within' the artefact boundary) and transferring it to the artefact in the other design concept. Ask yourself: "has the designer extracted elements from one concept and transferred them to, or mixed them up with, the other?"
- Relational combinations operate on 'whole' design artefacts and combine them through *external* relations (i.e., 'outside' the artefact boundary). Ask yourself: "has the taken both base concepts and related them together?". Note that a characteristic of relational combinations is that both base concepts are identifiable in more-or-less their original form, although some minor modifications might occur to facilitate the new relation.

There may be cases that you 'miss' instances of combination, i.e., where the participant has truly combined two concepts together but you are unable to infer that from the sketch. This is to be expected and is the reason for the 'NC' code. Please only use the F R or A codes if you can identify elements of both base concepts in the combined concept.

[Coding procedure on next page]

3.5. Coding procedure

The coding procedure guides you through the process of coding each design concept. Beginning with an uncoded concept, follow the flow chart in Figure and the accompanying list of instructions. Every path in the flowchart leads you to the application of a code (dark grey boxes) listed in Table .

The following steps for coding the concepts correspond to the grey circles in Figure .

1. Interpret the design concept created by the participant and the corresponding base concepts.
2. Ask yourself: has a design concept been created that is understandable and satisfies the design brief?
    a. If not, you will select one of the unsuccessful trial codes (I, U or O).
    b. If so, move on to 3.
3. Consider the things being represented in the design concept. Try to identify and mentally demarcate the design artefact (or artefacts) from the existing products and generic entities.
4. Ask yourself: has the new concept been created through combination of the base concepts? Specifically, can you identify elements of both base concepts in the combined concept?
    a. If not, select the non-combination code (NC).
    b. If so, select, a featural (F), relational (R), or ambiguous (A) combination.

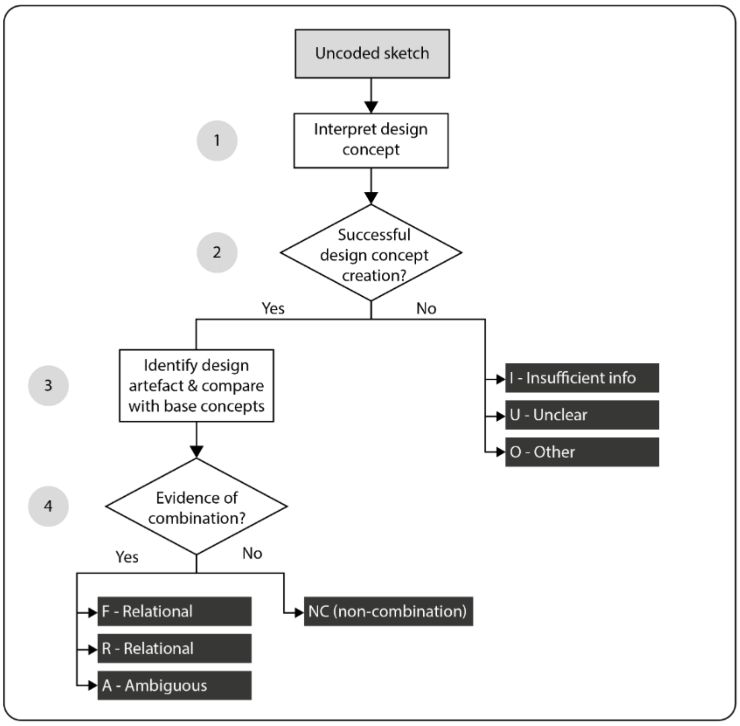


Figure 4 – Flow chart of coding procedure

4. Example

An example is provided below, showing the base concepts (Figure), and a featural (Figure), relational (Figure) and non-combination (Figure) response.

Figure shows a pair of design concepts created in response to the following design brief: ‘generate concepts for products that may facilitate water transportation in developing nations’.

- The concept one the left is described as ‘a water container bag, to make it easier to carry water’.
- The concept on the right is a water container with a strong elastic rope and hooks so that it may be attached to other objects.

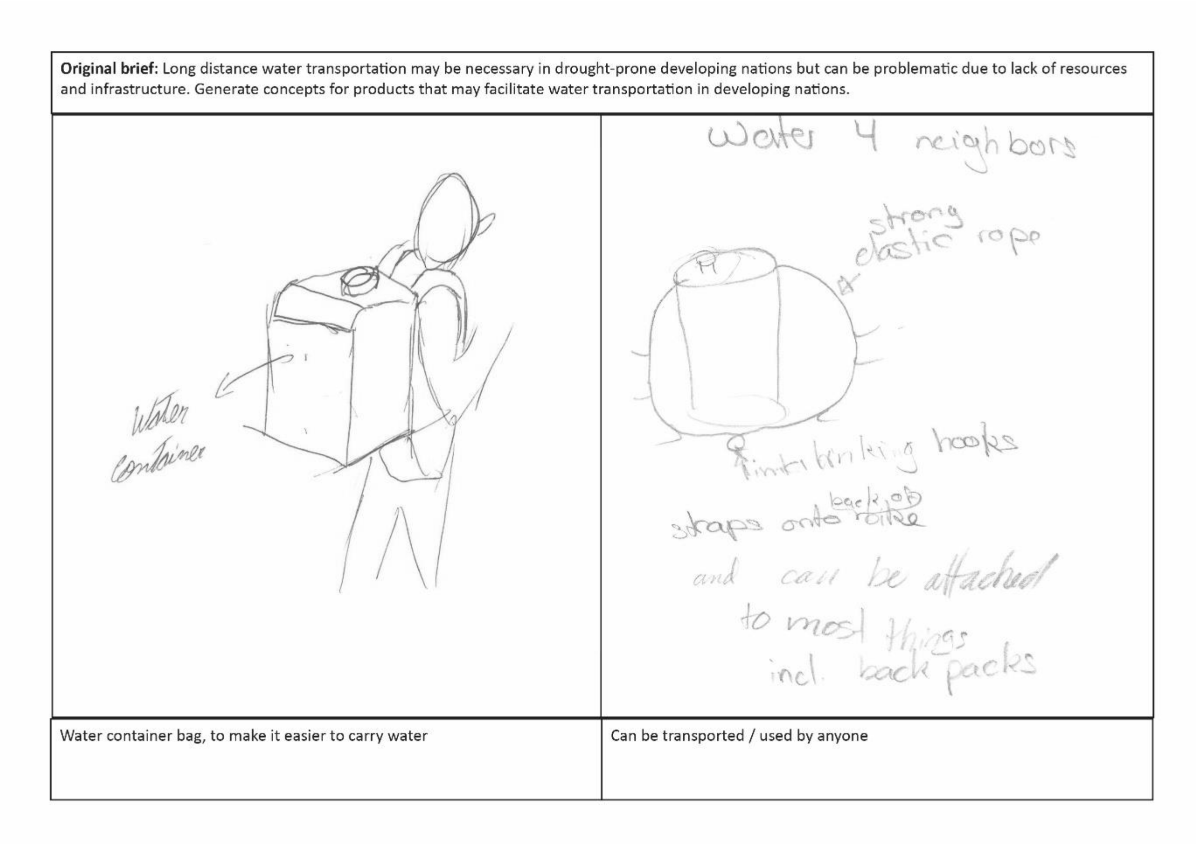


Figure 5 – An example of a pair of base concepts and design brief that they were created in response to**.**

Featural combination (Figure). In this example, the hooks from the right base concept have been transferred to the bag-carrier from the left concept to create a new design artefact. It is this transfer of a 'difference' that makes it a featural combination.

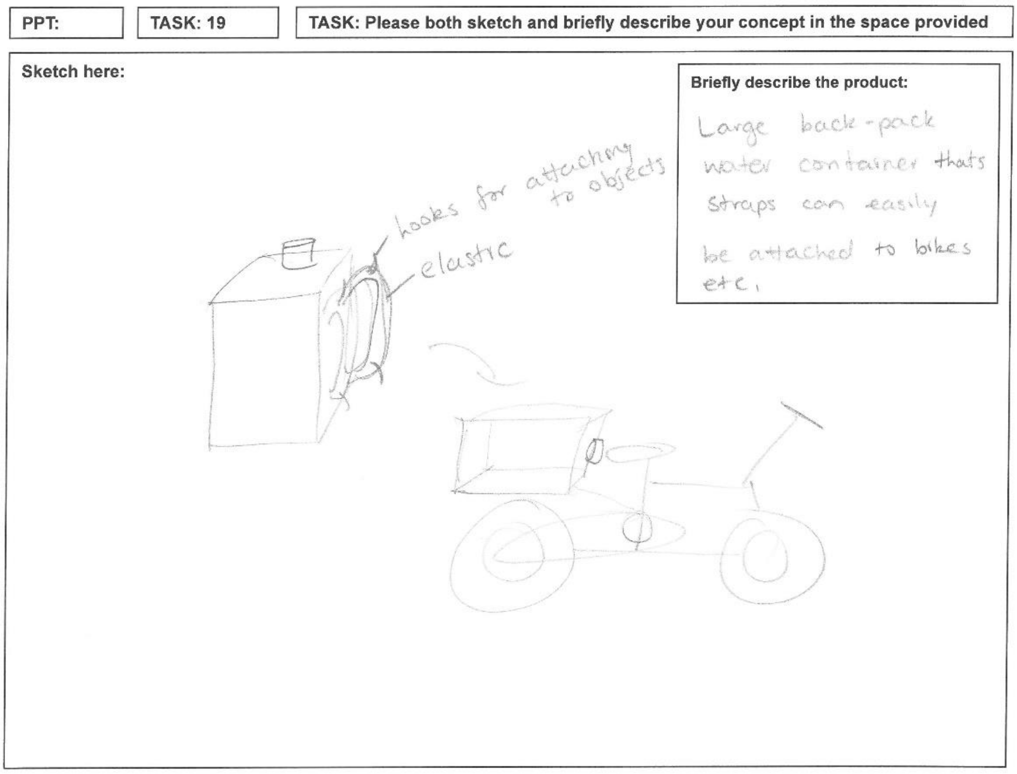


Figure 6 - An example of a featural combination

Relational combination (Figure). In this case, the artefacts from both base concepts can be seen in the new concept. The description of this concept is 'Water container can be attached to the backpack with the strong elastic rope'. This is a good example of a relational combination because it explicitly mentions the artefacts from both base concepts (water container, backpack) and a relation between them (water container – *attaches to* – backpack).

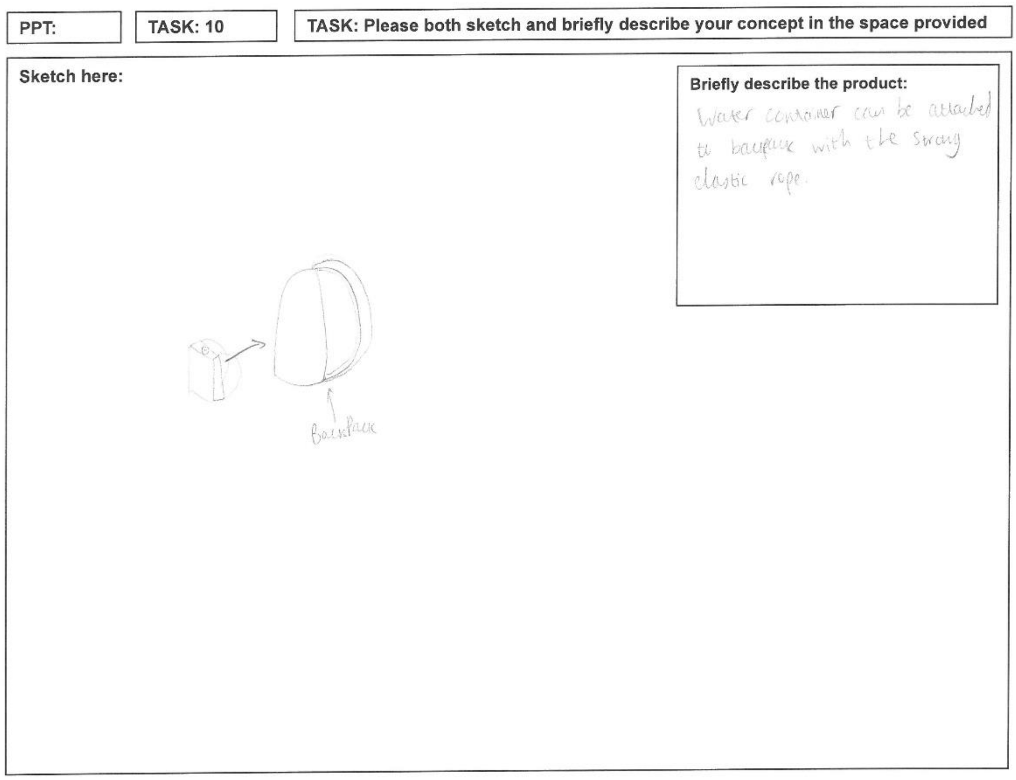


Figure 7 - Example of a relational combination

Non-combinational ideation (Figure). In this example, the artefact is a water carrier in the form of a suitcase. This addresses the brief and appears clear enough to be understood and thus it is a successful instance of ideation. The combined concept contains a commonality from both base concepts (they are all water containers). It is not clear, however, that any differences have been transferred from one concept to the other. **Thus, this cannot be coded as an example of combination.**

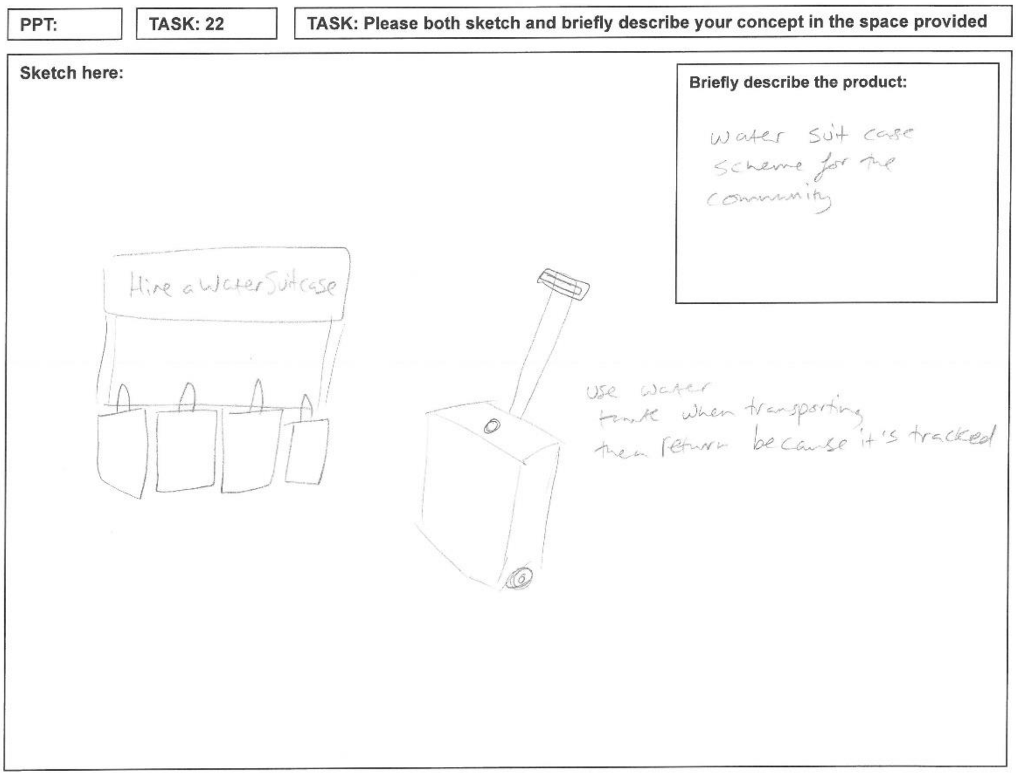


Figure 8 - Example of a non-combinational response

5. Conclusion

Use the materials, procedure and codes presented in Section 3 to code the design concepts that have been given to you.

APPENDIX 7 – SIMILARITY EXPERIMENT RESULTS

APPENDIX 7A – SIM-P1: OBSERVATION

Table A7-1 - Observations of Sim-P1: Issues and implications

Ref.	Issue	Implications
A	Similarity judgements of 16 concepts takes approximately 10-12 minutes	The duration of the experiment could be increased to include more trials (if desirable), without leading to excessive fatigue or boredom for the participants.
B	Some participants note that their initial trials are likely not accurate as it took some time to get used to the task	If participants change in how they respond to the task over time then this might be a source of noise in the data. Randomisation of the stimuli presentation order should mitigate any such noise to an extent. Participants may benefit from practise trials that are not analysed with the 'main' data.
C	The lack of written descriptions on some concepts is confusing / distracting. Some participants seem to rely heavily on the descriptions for their judgements. (The written descriptions in this case are those that were requested to be placed in the grey box on the sketch sheet. Some of the designers who created the design concepts used as stimuli did not provide such a description.)	The filtering criteria for Stimuli set A did not necessitate that all stimuli needed to have additional written descriptions. It was initially expected that participants would focus on the sketches and annotations. Increased homogenisation of the stimulus pairs in terms of their accompanying descriptions could mitigate potential confounds and would require relatively few resources.
D	Some participants offered insights into the mental process through which they made similarity judgements. One participant focused on the purpose of the concept as the primary determinant of similarity.	The model of mental concept representation in research phase 1 does not explicitly account for purpose and does not distinguish between function and purpose. The stimuli creation method may not be capturing all variables that contribute to an individual's perception of design concept similarity.
E	Three participants expressed difficulty with interpreting the sketches and attributed this difficulty to the use of a monitor. All three advocated for the presentation of stimuli on paper.	The initial decision to present the stimuli on the computer monitor was motivated by a desire to minimise the physical distance between stimuli and response boxes and reduce the need for back-and-forth head movements. This may have been an unnecessary consideration at the expense of the clarity of sketches and annotations.
F	Some participants expressed difficulty in establishing a consistent relative measure of similarity between pairs.	This may be attributable to: the relative size of the solution spaces for each design task, the relatively small number of comparisons in the experiment, or the lack of a design brief.

APPENDIX 7B – SIM-P2: LIST OF 'OTHER' CODES IN THE SIMILARITY EXPLANATION TASK

Table A7-2 - List of responses coded 'other' in Similarity Experiment 2.

Ref	Ptp	Pair	Explanation	Sim
			Assertion or negation of commonality / difference	
1	1	Sb_12	These are basically the exact same concept, each with more development of different aspects.	1
2	4	Sb_9	completely different methods	5
3	4	Sb_18	very different solutions to problem	4
4	4	Sb_29	different ideas	5
5	4	Sb_30	highly different ideas	5
6	4	Sb_40	different ideas	5
7	5	Sb_9	dint fully understand what concept 2 is doing ... but i dont see much overlap between the two concepts.	5
8	5	Sb_16	Some what similar in that the outcome is the same but process is slightly different.	3
9	5	Sb_19	completely different issues.	5
10	6	Sb_12	Both are essentially the same product; a trash compactor.	1
11	6	Sb_20	Entirely different problem being targeted and entirely different solutions.	5
12	7	Sb_2	same products, same scenario!	1
13	8	Sb_9	Very different in every way I can think of. Also can't really tell what either of them do exactly.	5
14	8	Sb_10	They have very different forms, and do very different tings.	5
15	8	Sb_36	Very similar idea, but products delivering it are very different.	3
16	9	Sb_4	Conceptually identical	2
17	9	Sb_11	Virtually identical	1
18	10	Sb_9	no connection what so ever	5
19	10	Sb_11	exact same concept	1
20	10	Sb_20	no connection what so ever	5
21	11	Sb_11	Same solution for the problem	1
22	11	Sb_19	The solutions are for different problems, hence completely different	5
23	11	Sb_39	The solution is built in the chair	5
			Ambiguous / potentially thematic	
24	4	Sb_39	both related to the office chair	5
25	5	Sb_11	very similar, use case could be developed in the first to meet the needs that the second has identified.	1
26	5	Sb_30	one transports one makes it.	5
27	11	Sb_29	One is a method of collecting the water while the other is a method of transportation	5

'Ref' is used to refer to the response in the main text. Each response includes participant number (ptp), rated similarity, pair number (Pair), the explanation for similarity and the numerical similarity rating.

APPENDIX 7C – SIM-P2: INTERVIEW RESPONSES. INTERPRETATION, IMPLICATIONS AND ACTIONS.

Table A7-3 - Interpretaton of resposnes from follow-up interview (Sim-P2)

Ptp	Issue	Methodological implication	Action
Q1 - Do you have any general thoughts on the experiment; any comments you would like to make?			
G1_PP1	Participant spontaneously opened a digital notepad and used it to copy / paste responses. The goal was to save time typing.	Potentially limits breadth of responses or naturalness of approach.	Instruct participants not to copy and paste.
	Participant provides an introspective report of a conceptual hierarchy they use for determining the similarity of pairs of design concepts.	N/A	N/A
G1_PP2	Participant expresses difficulty when asked for similarities of seemingly very different concepts and vice versa.	N/A	N/A
	Participant notes that some sketches were problematic in that they lacked information.	Sketch quality or degree of information may have been interfering with processing. Filtering may not have been thorough enough.	Conduct another round of filtering and exclusion for subsequent stimuli.
G1_PP3	Participant volunteers information about their own priority system for assessing the similarity of design concepts	N/A	N/A
G1_PP4	No response	No response	No response
G2_PP1	Participant found it difficult to discriminate between the task instructions for the similarity	Task instructions may not be clear. Alternatively, participants may find it difficult to discrimination between rationale for *similarity*	Check task instructions for clarity.

Ptp	Issue	Methodological implication	Action
	explanation task and the commonality and difference listing task.	and lists of *commonalities.* Although these are theoretically distinct, their meanings might be colloquially ambiguous.	Future research may want to avoid giving participants both a similarity explanation and a commonality and difference listing task. Participants may get confused.
G2_PP2	Participant notes that the second task (commonality and difference listing) makes them reflect on their responses to the first task (similarity explanation).	N/A	N/A
G2_PP3	Difficulty with quality of concept sketch, names one concept in particular	Sketch quality or degree of information may have been interfering with processing. Filtering may not have been thorough enough.	Conduct another round of filtering and exclusion for subsequent stimuli.
Q2 - Were there any elements of the experiment which you found to be particularly difficult?			
G1_PP1	Participant expresses preference for fewer concepts.	Too many concepts may lead to boredom or fatigue.	Reduce the number of stimuli in subsequent phases of research.
G1_PP2	Participant expresses difficulty in wording their responses.	Difficulty in providing responses may increase the duration of the study, thus leading to boredom or fatigue. Additionally, participants may provide low-effort responses.	Consider trialling verbal responses in subsequent phases or future research.
G1_PP3	Participant found it difficult to keep track of whether they were listing commonalities or differences. They also struggled to find meaningful differences to list for certain pairs.	Confusion over trial instructions could cause data collection errors.	Revise presentation format for lists of commonalities and differences. Encourage participants to tick off or otherwise indicate when they have completed a list for each stimuli.
G1_PP4	Participant highlights that listing the differences for highly similar concepts is challenging.	N/A	N/A

Ptp	Issue	Methodological implication	Action
G2_PP1	Hard to read some sketch annotations	Could potentially cause a number of issues: participant may pay less attention to annotations over time and start relying more heavily on typed description. Participant may miss key description of e.g. functional or aesthetic information in concept sketches if they cannot understand the annotation	Conduct another round of filtering and exclusion for subsequent stimuli.
	Participant found that they were capable of keeping four briefs in their head and swap between them and suggests that four design tasks is a manageable number.	N/A	N/A
G2_PP2	Keyboard issue	Potential problems: slows down participants responses by slowing typing speed, frustrates participant and causes distraction,	Replace keyboards.
G2_PP3	Participant noted that they ran out of time in a few instances, but not in others, presumably this means the second half of the experiment since that was the only one that was timed.	Participants may not list as many commonalities or differences as they can perceive. This could have negative consequences for the analysis.	Consider removing time limits and giving participants unlimited time for the listing task.
Q3 - How do you feel about the means of data entry (i.e. via laptop)			
G1_PP1	Participant would have preferred to hand write the responses because they make mistakes when typing	Keyboard entry could cause some participants difficulties with clearly communicating their responses.	Consider alternative input method in subsequent phases.
	When asked whether they would prefer to provide verbal responses, the participant said it could be easier for them, although they think it might have taken longer to get their point across clearly. The close by saying that they don't mind the typed input.	As above.	As above.

Ptp	Issue	Methodological implication	Action
G1_PP2	No response.	N/A	N/A
G1_PP3	Keyboard issue	As above	Replace keyboards.
G1_PP4	Preference for typing over writing	N/A	N/A
	Keyboard issue	As above	Replace keyboards.
G2_PP1	Preference for typing	N/A	N/A
G2_PP2	Preference for typing	N/A	N/A
G2_PP3	Preference for typing over writing	Preference for typing over writing	N/A
	Keyboard issue	As above	Replace keyboards.
Q4 - Do you feel that the accuracy of your similarity judgements was consistent throughout the experiment, or do you think it changed?			
G1_PP1	Participant advocates for a practise trial. They note that as they progressed through the study, prior judgements would form datums for subsequent judgements.	Accuracy of judgements may change over time.	Should be partially addressed by pseudo-randomisation. Consider including practice trials.
G1_PP2	Again, similarity judgements partially determined by prior examples. When prompted, they suggest that they began to develop a consistent system after aprx. 5 stimuli.	Accuracy of judgements may change over time.	Should be partially addressed by pseudo-randomisation. Consider including practice trials (5 or more).
G1_PP3	Participant started associating qualitative labels with numbers on the scale. For example, “4 and 6 are 'slightly more similar' and 'slightly less similar'” and “7 and 8 as 'really really similar, but not quite identical'”.	The rating scale provides qualitative labels at either end of the scale, but none in the middle. Participants may thus interpret the middle of the scale differently from one another.	Consider adding qualitative labels along the rating scale.

Ptp	Issue	Methodological implication	Action
G1_PP4	Participant developed some kind of mental system but cannot articulate the details of it	N/A	N/A
G2_PP1	Participant says they felt inclined to provide different justifications for similarity, they may have given more detail in an effort to differentiate previous answers	Participants answers could be skewed by some kind of expectation derived from the repetitive nature of the experiment.	Consider telling participant that repetition is fine (but not to copy and paste).
G2_PP2	Participant thinks their accuracy improved over the course of the experiment. They used prior judgements as references for subsequent judgements.	Accuracy of judgements may change over time.	Should be partially addressed by pseudo-randomisation. Consider including practice trials.
G2_PP3	Participant notes that their use of the scale changes over time. Suggests that some examples at the start might be useful.	Accuracy of judgements may change over time.	Should be partially addressed by pseudo-randomisation. Consider including practice trials.
Q5 - Do you feel that the manner in which you determined similarity changed throughout the study?			
G1_PP1	Participant refers to previous response (see Q4)	N/A	N/A
G1_PP2	Not clear		
G1_PP3	When prompted, participant suggests that their similarity ratings would differ for design concepts from different design tasks.	This is to be expected.	N/A

Ptp	Issue	Methodological implication	Action
G1_PP4	User mentions specific criteria: user cycle, user experience, where it is used, cost, capacity (specific to water example). They mention that "different issues" would make it easy to spot differences between concepts. Aesthetic similarity comes into play. May haven taken 10 to 15 concepts to get to "quite confident" about how different the concepts were	Participants may require a certain number of trials before arriving at consistent similarity ratings.	Consider adding practise trials (10-15 or more).
G2_PP1	Manner of responses "probably evolves" over the first 5 to 10. Examples of developing criteria: "did the same thing" "both biodegraded". Further evidence of prior concepts influencing new ones Certainly an evolution and perhaps different criteria per DT, but not conclusive "more consistent" after first 5to10	Participants may require a certain number of trials before arriving at consistent similarity ratings.	Consider adding practise trials (10-15 or more).
G2_PP2	The participant might notice a commonality (such as two design concepts involving desk chairs), but then reduce the relative importance of that commonality owing to the expectation that desk chairs would be present in a design task set in an office context.	The specific wording of the design brief might influence participant similarity ratings.	This justifies the inclusion of the design briefs alongside the stimuli (unlike Sim-P1 that only included the design concepts). Future research should include the relevant design task.
G2_PP3	Participant appears to give a rating, provide the explanation and then based on that explanation, change their rating.	Ratings are being adjusted based on the similarity explanations. This should be avoided if possible to stop the ratings being muddled with subsequent cognitive processing.	Ensure that similarity ratings cannot be altered after moving to a different task.

Ptp	Issue	Methodological implication	Action
Q6 - Do you feel that you had enough time to list the commonalities or differences of a pair of concepts?			
G1_PP1	No response	N/A	N/A
G1_PP2	No response	N/A	N/A
G1_PP3	1 minute perhaps not enough, 2 minutes would have been better.	Different amount of time for different tasks to be expected.	Consider removing time limit on listing task
G1_PP4	1 minute "just enough" Sketches with more annotations and sketch complexity increase time taken to list commonalities or differences.	If sketch complexity and volume of annotations influence time taken to list commonalities and differences, then it could skew the number listed, and thus interfere with our testing of the structural alignment predictions.	Conduct another round of filtering and exclusion for subsequent stimuli.
G2_PP1	Void	N/A	N/A
G2_PP2	1 minute sufficient Having been exposed to the sketches already may speed up responses in the second half.	Support for 1 minute being sufficient	N/A
G2_PP3	1 minute sufficient - some could go on listing more concepts "all day" (hyperbolically, endlessly)	Support for 1 minute being sufficient	N/A
Q7 - Please comment on your energy / effort levels towards the end of the experiment			
G1_PP1	Mild fatigue, not "exceptionally taxing"	N/A	N/A
G1_PP2	Effort, repetitiveness, drained afterwards	Excess fatigue could lead to low-effort responses.	Consider ways to reduce experiment length.
G1_PP3	Really long	Excess fatigue could lead to low-effort responses.	Consider ways to reduce experiment length.

Ptp	Issue	Methodological implication	Action
G1_PP4	Rushing to get to end	Excess fatigue could lead to low-effort responses.	Consider ways to reduce experiment length.
G2_PP1	Void	N/A	N/A
G2_PP2	Participant liked that the second section was timed, made it feel quicker than first section	N/A	N/A
G2_PP3	Participant had to will themselves to finish	Excess fatigue could lead to low-effort responses.	Consider ways to reduce experiment length.
Q8 - Do you have any suggestions for how the experiment might be improved?			
G1_PP1	Recommends to force a break at part 1. Input method and procedure fine. Would not prefer hand written	Repeat of previous issues about fatigue and data input.	N/A
G1_PP2	Suggests different box for each difference, but no real reason for it Talks about drop down boxes	N/A	N/A
G1_PP3	Suggests doing each design task sequentially in order to reduce the amount of switching back and forth required. Too random, too complex	Switching between different design tasks might be causing confusion.	The randomisation is necessary, but consider reducing the number of design tasks and total number of stimuli to make the study easier.
G1_PP4	No suggestions	N/A	N/A
G2_PP1	Comments on high number of concepts	Could be related to boredom / fatigue.	Consider reducing the number of design tasks and total number of stimuli to make the study easier.
G2_PP2	Suggests having commonalities or differences written on the sheet to avoid switching back and forth, reduce confusion	Presentation of instructions could be causing difficulties for participants.	Consider presentation of instructions for listing commonalities or differences.

APPENDIX 7D – SIM-P3: AGGREGATE LEVEL DATA

Table A7-4 - Aggregate level data for the Sim-P3 variables for the stimuli associated with DT03

	Sim(all)			Sim(com)			Sim(dif)			Com			Dif(tot)			AD		ND	
Pair	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	M	SD
1	8.37	0.808	35	8.24	0.903	17	8.53	0.743	15	4.53	1.841	17	2.07	1.387	15	1.53	1.356	0.47	0.834
2	7.47	0.950	32	7.54	1.127	13	7.35	0.862	17	4.00	2.082	13	2.41	1.121	17	1.76	1.091	0.59	0.712
3	6.94	1.197	33	7.08	1.038	13	7.07	1.280	15	2.77	1.423	13	3.73	2.017	15	2.80	1.897	0.87	1.246
4	6.00	1.766	35	6.06	1.482	16	6.00	2.171	15	4.00	1.265	16	2.27	1.033	15	1.47	1.187	0.87	0.834
5	3.80	1.922	35	3.63	1.857	16	3.80	2.111	15	3.44	1.365	16	3.73	1.580	15	2.67	1.047	0.93	0.961
6	6.82	1.381	34	6.87	1.552	15	7.07	1.223	15	4.00	2.299	15	3.47	1.187	15	2.27	1.100	1.33	0.976
7	3.14	1.537	35	3.00	1.837	17	3.08	1.188	13	3.82	1.944	17	3.08	1.498	13	2.08	1.382	0.46	0.660
8	2.51	1.401	35	2.67	1.589	15	2.47	1.356	15	1.80	1.082	15	3.53	1.922	15	2.00	1.000	1.73	1.831
9	1.59	0.756	32	1.77	0.927	13	1.50	0.632	16	2.15	0.899	13	4.13	2.335	16	3.06	1.692	0.81	1.167
10	1.44	0.613	34	1.36	0.497	14	1.53	0.717	17	1.79	1.188	14	4.59	2.093	17	2.82	1.510	1.94	2.277

Mean (M), standard deviation (SD) and number of responses (n) for each variable in DT03. The value of n for AD and ND is the same as for Dif(tot)

Table A7-5 - Aggregate level data for the Sim-P3 variables for the stimuli associated with DT06

	Sim(all)			Sim(com)			Sim(dif)			Com			Dif(tot)			AD		ND	
Pair	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	M	SD
11	8.44	0.613	34	8.46	0.519	13	8.47	0.640	15	3.77	1.423	13	2.07	1.534	15	1.80	1.320	0.27	0.594
12	8.54	0.611	35	8.62	0.650	13	8.47	0.514	17	4.92	1.891	13	2.24	1.091	17	1.35	1.169	0.88	0.857
13	7.68	0.912	34	7.86	0.770	14	7.79	0.802	14	4.29	1.899	14	3.21	1.626	14	1.79	0.975	1.50	1.345
14	5.60	2.003	35	5.40	1.844	15	5.31	2.287	13	3.73	1.580	15	4.69	1.797	13	2.23	1.363	2.46	1.984
15	2.71	1.426	34	2.65	1.498	17	2.93	1.387	15	2.88	1.364	17	3.00	1.000	15	2.00	1.195	1.00	0.756
16	6.57	1.335	35	6.64	1.646	14	6.50	0.894	16	4.07	1.900	14	2.19	1.223	16	0.69	0.704	1.50	1.265
17	3.49	1.442	35	3.79	1.626	14	3.50	1.314	12	2.86	1.834	14	3.58	1.782	12	2.00	1.477	1.50	1.168
18	2.58	1.226	33	2.50	1.211	16	2.60	1.352	15	2.38	1.408	16	4.13	2.386	15	2.27	1.534	2.00	1.927
19	1.30	0.585	33	1.33	0.651	12	1.14	0.363	14	2.42	1.379	12	3.50	1.698	14	2.21	1.424	0.93	0.616
20	1.06	0.236	35	1.07	0.267	14	1.08	0.277	13	1.71	0.914	14	5.54	2.727	13	2.31	1.377	3.23	2.920

Mean (M), standard deviation (SD) and number of responses (n) for each variable in DT06. The value of n for AD and ND is the same as for Dif(tot)

Table A7-6 - Aggregate level data for the Sim-P3 variables for the stimuli associated with DT07

	Sim(all)			**Sim(com)**			**Sim(dif)**			**Com**			**Dif(tot)**			**AD**		**ND**	
Pair	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	M	SD
21	7.71	0.799	34	7.81	0.834	16	7.50	0.760	14	5.13	1.821	16	3.57	1.399	14	3.29	1.139	0.29	0.611
22	7.74	1.221	35	7.59	1.121	17	7.81	1.377	16	4.12	1.965	17	2.44	1.413	16	1.56	1.459	0.88	0.806
23	6.09	1.485	34	5.81	1.721	16	6.21	1.188	14	3.94	1.289	16	2.86	1.512	14	2.21	1.528	0.64	1.151
24	4.26	1.704	35	4.29	1.773	14	4.07	1.580	15	3.79	2.007	14	3.47	0.915	15	2.00	1.309	1.60	1.454
25	3.29	1.582	35	3.12	1.654	17	3.38	1.586	16	3.82	2.186	17	3.88	1.857	16	3.19	1.328	0.69	0.873
26	5.00	1.645	35	5.20	1.699	15	4.83	1.689	18	4.60	1.549	15	4.44	2.332	18	3.11	1.676	1.33	1.609
27	2.91	1.463	35	3.12	1.616	17	2.86	1.406	14	3.12	1.364	17	4.14	1.351	14	3.43	1.399	0.71	0.914
28	1.97	0.951	33	1.93	1.033	15	2.00	0.935	17	2.87	1.767	15	5.53	3.676	17	2.88	1.867	2.71	2.443
29	1.60	0.775	35	1.57	0.852	14	1.73	0.799	15	1.64	0.842	14	3.93	1.792	15	3.07	1.580	0.80	1.146
30	1.54	0.886	35	1.63	0.957	16	1.40	0.828	15	2.00	1.461	16	5.20	3.189	15	2.27	2.187	3.00	2.726

Mean (M), standard deviation (SD) and number of responses (n) for each variable in DT06. The value of n for AD and ND is the same as for Dif(tot)

Appendix 8 – Combination results supplementary data

Appendix 8A – Visual representation of raw data for Combo-P3, grouped by design task

Figure A8- 1 - Square area chart showing the type of responses created by the participants in Combo-P3. Columns are grouped by design task.

APPENDIX 8B - AGGREGATE LEVEL DATA FOR COMBO-P3

Table A8-1 - Aggregate level data for Combo-P3

		Count of combination code						Difficulty ratings			Sim(all)		
DT	**Pair**	**F**	**R**	**A**	**Nc**	**O**	**n**	**M**	**SD**	**Mdn**	**M**	**SD**	**Mdn**
03	**1**	4	0	10	9	3	26	2.27	*1.28*	2.0	8.37	*0.81*	9.00
	2	10	1	0	9	6	26	2.96	*1.48*	2.5	7.47	*0.95*	8.00
	3	11	7	0	5	2	25	2.68	*1.18*	2.0	6.94	*1.20*	7.00
	4	17	5	0	5	0	27	2.11	*1.19*	2.0	6.00	*1.77*	6.00
	5	12	5	0	7	1	25	2.64	*1.19*	2.0	3.80	*1.92*	4.00
	6	11	0	0	13	3	27	2.30	*1.30*	2.0	6.82	*1.38*	7.00
	7	21	2	0	3	1	27	2.44	*1.15*	2.0	3.14	*1.54*	3.00
	8	16	3	1	3	3	26	3.00	*0.96*	3.0	2.51	*1.40*	2.00
	9	17	4	1	3	2	27	3.07	*1.44*	3.0	1.59	*0.76*	1.00
	10	10	14	2	0	1	27	2.78	*1.25*	3.0	1.44	*0.61*	1.00
06	**11**	4	0	0	21	2	27	2.48	*1.37*	2.0	8.44	*0.61*	8.50
	12	10	0	1	13	0	24	2.63	*1.13*	2.5	8.54	*0.61*	9.00
	13	16	0	0	8	1	25	2.72	*1.43*	2.0	7.68	*0.91*	8.00
	14	7	5	8	3	1	24	2.48	*1.41*	3.0	5.60	*2.00*	6.00
	15	1	22	1	1	3	28	2.57	*1.23*	2.0	2.71	*1.43*	2.00
	16	15	1	0	5	4	25	2.96	*1.27*	3.0	6.57	*1.33*	7.00
	17	13	8	0	0	1	22	2.95	*1.40*	3.0	3.49	*1.44*	3.00
	18	13	9	0	2	4	28	2.86	*1.11*	3.0	2.58	*1.23*	3.00
	19	4	18	0	1	3	26	3.04	*1.31*	3.0	1.30	*0.59*	1.00
	20	13	4	0	2	6	25	3.96	*1.24*	4.0	1.06	*0.24*	1.00
07	**21**	15	1	0	9	1	26	2.42	*1.10*	2.0	7.71	*0.80*	8.00
	22	9	2	1	12	3	27	3.26	*1.32*	3.0	7.74	*1.22*	8.00
	23	13	4	2	4	3	26	2.52	*1.16*	2.0	6.09	*1.48*	6.00
	24	22	1	0	2	1	26	2.42	*1.03*	2.0	4.26	*1.70*	5.00
	25	4	17	1	2	2	26	2.35	*0.98*	2.0	3.29	*1.58*	3.00
	26	4	3	17	3	0	27	1.85	*0.86*	2.0	5.00	*1.64*	5.00
	27	14	3	8	1	1	27	3.00	*1.21*	3.0	2.91	*1.46*	3.00
	28	18	2	0	1	7	28	3.71	*1.01*	4.0	1.97	*0.95*	2.00
	29	1	1	19	3	2	26	2.35	*0.94*	2.0	1.60	*0.77*	1.00
	30	7	5	0	3	11	26	3.65	*1.41*	4.0	1.54	*0.89*	1.00

Note: Includes counts of combination codes and summary data for difficulty ratings and rated similarity. Similarity ratings are taken from Sim-P3. F featural combinations, R relational combinations, A ambiguous combinations, O other responses, n number of responses for the given stimulus, M mean, SD standard deviation, Mdn median.