View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

Computer-based 'mental set' tasks: An alternative approach to studying design fixation

Maria Adriana Neroni, first author
Department of Engineering
University of Cambridge
Trumpington Street, Cambridge, CB2 1PZ
United Kingdom
man52@cam.ac.uk

Luis A. Vasconcelos, second author
Department of Engineering
University of Cambridge
Trumpington Street, Cambridge, CB2 1PZ
United Kingdom
lal40@cam.ac.uk

Nathan Crilly, third author
Department of Engineering
University of Cambridge
Trumpington Street, Cambridge, CB2 1PZ
United Kingdom
nc266@cam.ac.uk

ABSTRACT

The term 'design fixation' refers to a phenomenon where designers unknowingly limit the space within which they search for solutions. In an attempt to study this phenomenon experimentally, researchers typically set participants open-ended design problems, prime them with an example solution and measure their performance through a variety of subjective metrics. This approach gives rise to various problems, including limited data capture and highly subjective evaluation of design behavior. To address these problems, we studied design fixation with a computer-based task inspired by psychological paradigms used to study 'mental set' (also known as the 'Einstellung effect'). The task consisted of a game-like activity requiring participants to design a bridge within a specified budget. The use of a digital environment facilitated *continuous data capture during the design activities. The constrained task (and direct quantitative measures) permitted a more objective analysis of design performance, including the occurrence of fixation. The results showed that participants who developed a mental set during the task failed to find alternative, more efficient solutions in trials admitting multiple solutions, compared to the participants who did not fall victim to this mental block. In addition, during the process of designing, the occurrence of mental set resulted in participants adopting a less efficient design behavior and reporting a different subjective experience of the task. The method used and the results obtained show an exciting alternative for studying design fixation experimentally and promote a wider exploration of the variety of design activities in which fixation might occur.*

Keywords: *Design creativity, Design fixation, Mental set, Einstellung effect, Computer-based task, Design process*

1 INTRODUCTION

Designers of all disciplines – working on many different kinds of projects – are required to be creative if they are to arrive at new and useful solutions to the problems that they address. Design processes, whether systematic or intuitive, are often claimed to unlock this creativity by inspiring designers to consider a broad range of possible solutions and by discouraging premature commitment to any one of those solutions. Despite this, designers do become ‘set’, ‘blinkered’ or ‘blinded’ when developing ideas, limiting the way in which they interpret problems and explore possible solutions. In particular, potentially useful sources of inspiration or information can have the negative effect of restricting rather than freeing the designers’ imagination. This phenomenon is defined as ‘Design Fixation’ and refers to “situations where designers limit their creative output because of an overreliance on the features of preexisting designs, or more generally, an overreliance on a specific body of knowledge directly associated with a problem” [2 (p. 1)] (for narrower interpretations of design fixation see [1], for broader interpretations see [3,4]). Fixation is reported as a problem in professional design practice [5, 6] and has been measured and manipulated in experimental studies, many of which use engineering design problems [7, 8, 9]. Studies of design fixation are now sufficient in number that literature reviews have recently been published, focusing on the experimental methods used [10] and the findings obtained [11].

Research into design fixation has generated insights into why fixation occurs and how it might be mitigated. This is generally important for developing our understanding of design cognition, and particularly important for understanding the relationships between information, inspiration and creativity. However, like other aspects of human behavior, the knowledge we have about design fixation is dependent on the quality and variety of the methods we use to study it [12]. From that perspective, the existing studies of design fixation are quite homogenous in the experimental approach adopted (there is little methodological diversity) and that approach has some specific features and limitations. In particular, the open-ended nature of the design task that participants are given and the wide-ranging responses that result give rise to a complicated and subjective analysis of fixation. Alternative experimental approaches should be considered, and these might be applicable not just to the early-stage design activities that have been studied to date, but to other types of design activity also.

To address methodological issues with the established approach, we propose an alternative: studying design fixation with a computer-based task inspired by psychological paradigms used to study mental set [13, 14]. Doing this allows us to explore fixation effects as they might occur in later stages of the design process (once the overall concept is defined), thus complementing previous research on fixation in early-stage ideation activities. In addition, the use of a digital platform allows better data capture (because the design work is all digital), less variability in participants’ performance (because the design task is partially constrained and numerically tested) and a more objective analysis of design behavior (for the same reasons). The method used and the results obtained are reported here to offer a viable alternative for studying design fixation experimentally. We also hope that this will promote a wider exploration of the possible ways in which fixation might be studied and a greater appreciation of the variety of design activities in which fixation might be manifest.

2 LITERATURE REVIEW

2.1 Design Fixation Experiments

In their original study, Jansson and Smith [7] ran a number of experiments where designers, working individually, had to generate ideas in response to different problems (i.e. design a car-mounted bicycle rack, a measuring cup for the blind and a disposable spill-proof coffee cup). Alongside the design briefs, some of the participants were also presented with pictures of existing solutions. Jansson and Smith identified the occurrence of fixation when designers exposed to those pictures tended to repeat key features of the solutions that were represented. This behavior persisted even when participants received instructions to avoid repeating particular features of those example solutions. As these features were intentionally problematic (e.g. they contradicted the brief) this feature repetition was taken to be inadvertent and counterproductive. Since 1991, the basic approach taken in Jansson and Smith’s study has been adopted by many other researchers who set similar tasks and measure design performance with metrics that are either relatively objective (e.g. the number of final concepts, the number of different types of concepts and the repetition of key features from the example) or highly subjective (e.g. novelty, feasibility, relevance) [10]. What is often observed is that exposure to example solutions is associated with a reduction in the variety and quantity of solutions that designers generate in response to the design task (for a review of the results, see [11]).

In following Jansson and Smith's example, the vast majority of fixation studies seek to identify and understand the occurrence of design fixation in the early stages of a design process, especially during 'early ideation' or 'idea generation'. In these studies, researchers typically set participants ill-defined and ill-structured design problems (see [15]), which in turn permit a multitude of possible solution types. These particular features of the design activities that are studied have a number of consequences for fixation research methodology. In particular, the variety of possible solution types means that performance in these tasks must be evaluated with subjective metrics. Many studies involve a panel of experts scoring the design ideas in terms of the novelty, feasibility and relevance of the design outcomes that participants generate. However, variation in how the metrics are interpreted and applied gives rise to problems when trying to compare the design performance of participants within and between studies. Also, because the open-ended design problems that are set often lack the clear objectives and constraints that are typical of much design work [16], it is difficult to establish if unexplored areas of the design space have been inadvertently overlooked by participants or whether they have avoided them intentionally.

Beyond the methodological consequences outlined above, we might also question whether the open-ended early-stage ideation problems typically used in design fixation studies are the most important (or the only important) problem type addressed in design practice. Designers also engage with well-defined problems, and with late-stage design tasks. These more constrained design activities still demand creativity [17, 18], and might lend themselves to a more objective study of fixation. For example, many design tasks only permit the exploration of a well-bounded solution space (e.g. redesign of an electrical circuit, or design of a truss structure) where the challenge lies in creatively designing within those boundaries, rather than freely reframing the whole problem or context. Studying such tasks might permit better comparison between designers’ work because the outputs that they generate would be within the same category or domain. This would also permit more objective assessment of the outputs because the relationship between ideas might be directly quantified and those ideas could be subjected to consistent performance measures (e.g. using circuit simulators, or structural analysis).

In summary, the predominant focus of design fixation research on early stages of the design process has resulted in various methodological consequences related to the task that is set, the data collected and the ways in which that data is analyzed. In exploring alternative possible methods to study design fixation, the type of design activity focused on might change also.

Design fixation research originally took psychological research on mental blocks as useful precedents [1 (p. 4)], and we can look to those areas of research for alternative approaches. In particular, in the first studies of design fixation (and many of those that followed), research into two different psychological phenomena were cited as precedent: studies of ‘functional fixedness’ [19, 20] and studies of ‘mental set’ [14, 21, 22, 23]. The concept of design fixation perhaps inherited its name from functional fixedness, but the effect of interest is really more like mental set. This is because design fixation is not normally studied as a block to seeing new possible functions in existing structures (functional fixedness), but as a block in imagining new structures for specified functions (especially where an existing structure is known). Considered broadly, these 'new structures' might not just be products that realize a function, but also solutions to a problem or the approach, process or method by which those solutions are determined. A block to seeing new approaches is explained by mental set, and as such, experimental paradigms for studying mental set provide a valuable starting point for studying design fixation.

2.2 Mental Set Paradigms

Mental set refers to the development of a mechanized state of mind (also called the 'Einstellung effect'), which occurs when previous experience with a successful problem solving approach prevents alternative approaches being considered (even when those other approaches might be superior). The mental set phenomenon was originally studied by Luchins [14] with his water-jar task, sometimes referred to as the 'three jars' studies. In Luchins' studies, participants are presented with a series of arithmetic problems that are posed in the form of a puzzle involving volumes of liquid in different containers. There are three jars of different sizes and another larger container; participants are required to measure out the right amount of water to fill the container by adding and subtracting the volumes of the three jars through a series of 'pouring operations'.

In his studies, Luchins typically presented participants with ten problems. The first five problems could all be solved by the same complex sequence of pouring operations, even though the numbers involved varied (i.e. the volume of the three jars and of the container). These were termed ‘Einstellung’ problems (German for ‘mind set’) because the common solution method used for all of them was expected to induce a mental set for solving similar problems. The next two problems, termed ‘critical’, could be solved using the same method but a simpler (and easier) solution was also possible. Despite the simplicity of the alternative solution, the vast majority of participants persisted with the sequence they had used in the introductory problems. Then, Luchins posed an ‘extinction’ problem, which could not be solved using the original method but could be solved with the simpler method. Many participants said it was insoluble. A comparison group who were given only the extinction problem solved it quickly, showing that the problem was not intrinsically difficult, but was just made difficult by participants in the experimental group becoming ‘mentally set’ in their ways. Thus, rather than improving their performance, the experimental group’s additional experience of the general class of problem (arithmetic, pouring) blinded them to a simple solution which was found by almost everyone who had not had that additional experience [14, 21, 22, 23]. However, when Luchins presented the experimental group with two additional ‘critical’ problems (after the ‘extinction’ problem), he observed a slight increase in the participants’ tendency to use the alternative, simpler method. Thus, the ‘extinction’ problem helped participants to recover from the Einstellung effect.

Luchins’ striking demonstration that experience can induce inflexibility of thought has been successfully repeated using a variety of tasks, including those involving mathematical problems [24], verbal reasoning [25] and more visual tasks [26]. In addition, mental set has been shown both to affect people facing problems that are new to them

[27, 28, 29] and also experts tackling problems that are within their field of expertise [13, 30, 31, 32]. In particular, Bilalić et al. [13] conducted a series of studies documenting the existence of mental set in chess players. In these studies, researchers asked chess players to solve chess problems by finding the quickest way to achieve checkmate. Chess players were presented with a 2-solution problem, that could be solved using a more familiar five-move solution or a less familiar three-move solution, and were also presented with a 1-solution problem (‘extinction’ problem), that could be solved only using the less familiar three-move solution. Even expert chess players failed to find the optimal solution to the 2-solution problem when the familiar five-move solution was possible (but they were able to identify it when the five-move solution was not possible). This behavior was explained with reference to the phenomenon of mental set, or as Bilalić et al. said, because “good thoughts block better ones”.

Luchins’ water-jar task, Bilalić et al.’s chess problem and other experimental paradigms that have been developed to study mental set have a number of features that distinguish them from the methods typically used to study design fixation. In mental set paradigms, the solution space is restricted and some solutions are objectively preferable to others, either because they are correct [14], quicker or more efficient [13]. Adopting a mental set approach to design fixation offers the possibility to study design fixation in the exploration of a bounded solution space, where the objectives and constraints are well defined and where the relative merits of different solutions can be quantified directly. Mental set paradigms also offer the opportunity for the fixating solution to be one which the participants arrive at spontaneously (e.g. their first solution), rather than one which is directly provided by researchers (e.g. an example solution). Finally, mental set paradigms offer the opportunity to observe how fixation effects occur during the design process, rather than just inferring those effects from the design solutions that are generated (e.g. see Bilalić et al. [13] for a time-based analysis of chess players’ behavior).

3 THE PRESENT STUDY

With the objective of developing a mental set paradigm for studying design fixation, we developed a computer-based design task that would address some of the limitations identified in section 2.1 and yield some of the advantages described in section 2.2. We adapted a computer game platform to provide a digital environment in which bridge structures could be designed (with a budget constraint) and tested (with a specified load). This permitted a bounded exploration of the design space, the opportunity to test designs in situ and an objective measure of design performance.

By having participants complete a number of trials, we were able to structure our experiment according to the ‘three jars’ task used by Luchins [14]. In particular, we observed the performance of an experimental group, whose previous experience with a single-approach to the design task was expected to block their recognition of alternative approaches. The performance of this group was compared to that of a control group, whose experience with different approaches was not expected to prevent the occurrence of mental set.

4 METHOD

4.1 Participants

Forty participants (11 women) with an engineering qualification were recruited into the study by responding to posted advertisements; they were all drawn from the Department of Engineering of the University of Cambridge. Their average age was 25.19 (*SD* = 3.66) and they had an average of 5.97 (*SD* = 2.60) years of university education. The data from four outlier participants was later removed from the analysis because their performance (total time taken to complete the task) varied by two or more standard deviations from the mean of the group. The final sample therefore consisted of 36 participants.

Before starting with the task, participants were asked to indicate how many hours per month they played computer games (i.e. never, less than 24 hours, between 24 and 48 hours, between 48 and 72 hours, more than 72 hours). The majority of the participants (24 out of 36) stated that they spent less than 24 hours per month playing computer games. Only 2 participants said they played computer games more than 72 hours per month.

All participants gave informed written consent prior to the commencement of the study and received a small honorarium for their participation. The study procedures were approved by the local ethical review committee.

4.2 Materials and Procedure

The platform for the experimental task was an adapted version of the computer game ‘Pontifex’ [33], in which players must design a truss bridge that spans a river.

Participants were required to design a series of bridges that could each support a specified load using the available budget. The game included two modes. The ‘Design’ mode allowed participants to plan the bridge structure by arranging and modifying the structural elements, considering the required span (to cross a river) and the ‘anchor points’ that could be used to support the bridge (e.g. the banks of the river, or a rock at mid-span). The ‘Test’ mode allowed participants to load the bridge (by driving a train across it) and assess its performance (see Figure 1). To capture the participants’ activities during the game and the utterances they made, the ‘Fraps’ screen recording software was used [34]. This recorded every design and test cycle, including the placement, deletion and resizing of the structural elements, as well as the resulting cost of the bridge and its performance when subjected to loading.

[Figure 1]

The task consisted of a series of trials varying in the length of the gap that the bridges were required to span, and the number and location of the anchor points that could be used to connect the various structural elements. In each trial, the participants were required (and incentivized) to design the lowest cost bridge that would support the load. In particular, there were ten trials:

- five *single-approach trials*, in which successful bridges could only be designed by using all the available anchor points (otherwise the bridges would fail to support the load);
- four *dual-approach trials*, in which successful bridges could either be designed by using all the available anchor points or by not using all the anchor points (both types of

design could be within budget and support the load, but those designs that used all the anchor points would cost more to construct than was necessary);

- one *‘extinction’ trial*, in which successful bridges could only be designed by not using all the available anchor points (otherwise the budget would be exceeded).

Importantly, each of the ‘approaches’ that the participants could adopt would result in a set (or ‘family’) of design solutions for each trial, with different numbers and sizes of elements and different associated costs. Hereafter, we only distinguish between the different design approaches adopted and the broad sets of design solutions associated with those approaches. We do not distinguish between the different detailed geometries constructed or the resulting small differences in cost; the experimental trials were not set up to differentiate on that basis. However, to account for the variety of possible designs, in almost all the trials (except the extinction trial) the available budget was at least double the budget required to construct the optimal solution. This allowed participants to design any bridge structure that they were likely to attempt, regardless of the cost.

Participants were assigned to two experimental groups, which only varied in the order of the trials (see Table 1). In the experimental group, the trial order resembled the order of the jars problems in Luchins’ study [14]. Participants completed five single-approach trials, then two dual-approach trials, then an extinction trial, and finally two additional dual-approach trials. In the control group, participants received the same set of trials in a different order. Crucially, they started the task by completing the extinction trial and the single-approach trials were not grouped together. This experimental manipulation was expected to avoid the development of a mechanized state of mind because the initial trial did not require the use of all the anchor points but some subsequent trials did.

Each participant was tested individually in a private meeting room, seated at a computer screen. An initial introduction explained that they would be playing a computer game and answering some questions about their performance. Before starting with the experimental task, participants received a brief training session in order to familiarize themselves with the game platform. Their task in the training session was to design and test a tower, which familiarized them with both the Design and the Test modes. Once the participants were able to interact with the game platform without further assistance from the experimenter, they were provided with the written instructions concerning the experimental task. The overall aim of the task was to design the least expensive bridge possible that could withstand the load of a train moving across it without suffering any damage (i.e. broken structural elements or connections). To further motivate participants, they were told that their final score (calculated as the amount of money they had left after the ten trials, i.e. the unspent budget) would determine their ranking. At the end of the study, the highest ranked participant would receive an online shopping voucher. Additionally, to reduce the influence of participants' background and/or previous experience with similar design-activities, they were told that they would be constructing the bridges in an imaginary world (e.g., anchor points could be in mid-air, and not just on the ground). In all the trials, participants could iterate freely between the Design and Test modes in their attempts to design the least expensive bridge possible with each iteration later being counted as one ‘design-test cycle’. There was no time constraint to complete the task.

[Table 1]

To provide a more contextually rich account of the participants’ design behavior, they underwent a post-experimental interview (after playing the game) aimed at evaluating whether participants in the two groups had a different subjective experience of the task. In particular, participants were asked to rate on a 5-points scale (a) how much they enjoyed playing the game (1 = I did not enjoy it, 5 = I enjoyed it a lot), and (b) how difficult they found it to play the game (1 = it was extremely difficult, 5 = it was not difficult at all). Additionally, participants were required to indicate whether they had constructed the bridges in an automatic way or if they had thought about the different possibilities. Finally, they were asked to guess the aim of the study. At the end of the experiment, all participants received a quick debriefing explaining the main purpose of the study. The total testing time was about one hour per participant.

5 RESULTS

When working on the dual-approach trials, 15 out of 18 (83%) participants in the experimental group failed to design less expensive bridges; they instead continued to design following the approach used in the previous single-approach trials (see Table 2). This is similar to what Luchins found, and just as in his studies [14, 21, 22, 23], we observed that this tendency to automatically repeat the previously learned approach slightly decreased in the two dual-approach trials following the extinction trial (in the experimental condition two dual-approach trials preceded the extinction trial and two followed it). In particular, a Chi-square test with the proportion of dual-approach trials solved with the more expensive approach as the dependent variable revealed that the more expensive approach was used more often to design bridges in the dual-approach trials preceding the extinction trial (i.e. 21 out of 26 trials, or 81%) than in the dual-approach trials following the extinction trial (11 out of 26 trials, or 42%), $\chi^2(1) = 8.12$, $p = .004$, $\varphi = .39$. (For this analysis we removed 5 participants who were able to solve both the dual-approach trials preceding the extinction trial using the alternative, less expensive approach.)

5.1 Effect of mental set on the design solution

To quantify the effect of mental set on the participants’ design solution, we compared the experimental and the control group in terms of (a) the number of participants who designed the bridges using the more expensive approach in the dual-approach trials (even though an alternative, less expensive approach was possible), and (b) the proportion of dual-approach trials solved following the more expensive approach. A Chi-Square test with the proportion of participants who solved the dual-approach trials with the more expensive design approach as the dependent variable showed a significant difference between the two groups, $\chi^2(1) = 5.90$, $p = .02$, $\varphi = .40$, with more participants in the experimental group designing the bridges with the more expensive approach than participants in the control group (see Table 2). Similarly, a Chi-Square test with the proportion of dual-approach trials solved with the more expensive approach as the dependent variable revealed a significant difference between the two groups, $\chi^2(1) = 13.93$, $p = .000$, $\varphi = .31$, with participants in the experimental group more often adopting the more expensive approach than participants in the control group (see Table 2). Finally, an independent-samples t-test with the average cost of the bridges designed by the two groups as the dependent variable showed that the experimental group (*M* = \$30,077.39, *SD* = 2,701.22, *SE* = 636.68) designed more expensive bridges overall compared to the control group (*M* = \$27,461.11, *SD* = 1,999.76, *SE* = 471.35), $t(34) = 3.3$, $p = .002$, $d = 1.10$.

[Table 2]

5.2 Effect of mental set on the design process

To evaluate the extent to which developing a mental set influenced the participants’ design process, we compared the experimental group and the control group with regards to (a) the total time taken to design the bridges, (b) the number of structural elements used through the various design-test cycles, (c) the number of design-test cycles, (d) the cost of constructing the bridges in all the trials, and (e) the behavior exhibited when attempting to construct the least expensive bridge possible.

An independent-samples t-test with the average time that participants spent to solve all the trials as the dependent variable showed a significant difference between the two groups, $t(34) = 2.10$, $p = .04$, $d = .70$, with the experimental group spending more time overall to design the bridges compared to the control group (see Table 3). Similarly, an independent-samples t-test with the average number of structural elements used by participants in the two groups as the dependent variable showed that those in the experimental group added and subtracted significantly more structural elements during their design iterations compared to the participants in the control group, $t(34) = 2.57$, $p =$

.01, d = .85 (see Table 3). In contrast, an independent-samples t-test with the average number of bridge designs tested as the dependent variable showed no significant difference between the experimental and the control group, $t(34) = .76$, $p = .45$, $d = .35$ (see Table 3).

[Table 3]

Looking at the participants’ behavior allowed us to make observations about the ways in which they were designing, not just the design outcomes they arrived at. In particular, when comparing the first design-test cycles (across all trials) run by the participants in the two groups by means of an independent-samples t-test, we observed that the first designs tested by participants in the experimental group were significantly more expensive overall than the first designs tested by participants in the control group (experimental group: M = $31,059.26, SD = 2684.13, SE = 632.65; control group: M = $28,360.76, SD = 2394.04, SE = 564.28), $t(34) = 3.18$, $p = .003$, $d = 1.06$. Crucially, as shown in Figure 2, the cost of the first designs tested by the control group was also lower than the cost of the least expensive design solution, which demonstrates how participants in the control group often started with very low cost solutions, even though those solutions were not functional.

[Figure 2]

Following this observation, we compared the cost of each tested bridge design with the cost of the least expensive design solution (see Table 1). This allowed us to classify the participants’ design behavior:

- *‘stop if it works’*– designed and tested only one bridge without any further design-test cycles (design cost ≥ least expensive solution);
- *‘strong then cheap’*– started with a functional design that was more expensive than the minimum cost solution, and then reduced costs by eliminating structural elements through one or more design-test cycles (initial design costs > least expensive solution);
- *‘cheap then strong’*– started with a low cost design that did not support the load, and then strengthened the structure by incorporating additional structural elements through one or more design-test cycles (initial design cost < least expensive solution).

While the experimental group and the control group did not differ with regards to the frequency with which they exhibited the ‘stop if it works’ behavior (see Figure 3), opposite results were found for the ‘strong then cheap’ and ‘cheap then strong’ behaviors. In particular, a Chi Square test with the frequency of use of the ‘cheap then strong’ behavior as the dependent variables revealed a significant difference between the two groups, $\chi^2(1) = 15.33$, $p = .000$, $\varphi = .32$, with the control group more frequently exhibiting the ‘cheap then strong’ behavior than the experimental group. An additional one-way Analysis of Variance (ANOVA) with ‘type of behavior’ as the independent variable and average amount of money spent to design the bridges as the dependent variable was conducted to evaluate the extent to which adopting different types of design behavior influenced the amount of the available budget participants used to design the bridges. The use of the three types of behavior significantly affected the cost of the designed bridges, $F(2, 274) = 5.12$, $p = .007$, $\eta_p^2 = .03$. Planned Student's t-tests with the average cost of the bridges designed by participants as the dependent variable were conducted in order to evaluate the effect of each type of behavior on the amount of budget spent. Although the comparisons were planned prior to the experiment, the Bonferroni correction (0.05÷3 comparisons, alpha level to 0.017) was nevertheless implemented to minimize the probability of type 1 error, thus providing results that were more conservative. The results revealed that, compared to the 'strong then cheap' behavior (*M* = $34,234.94, *SD* = 10,701.35, *SE* = 1,261.17), adoption of the 'cheap then strong' behavior (*M* = $29,291.83, *SD* = 6,650.97, *SE* = 778.44) led participants to design less expensive bridges overall, $t(143) = -3.35$, $p = .001$, $d = .55$. No differences were found between the 'stop if it works' behavior (*M* = $30,967.71, *SD* = 10,157.41, *SE* = 884.09) and the 'strong then cheap' behavior, $t(202) = -2.15$, $p = .03$, $d = .31$, or between the 'stop if it works' behavior and the 'cheap then strong' behavior, $t(203) = 1.26$, $p = .21$, $d = .19$.

[Figure 3]

In line with previous studies on mental set [14, 15], we also compared the performance of the experimental group and the control group on the extinction trial. This allowed us to establish in what ways the occurrence of mental set influenced the performance of participants in the experimental group when working on this trial (for which the only approach that would work was different to the approach needed to solve the preceding single-approach trials).

All the participants were able to complete the extinction trial. However, an independent-samples t-test with the average number of structural elements used by participants in the two groups as the dependent variable revealed that those in the experimental group added and subtracted significantly more structural elements during their design iterations (*M* = 19.23, *SD* = 19.76, *SE* = 5.48) compared to the participants in the control group (*M* = 8.28, *SD* = 7.43, *SE* = 1.75), *t*(34) = 2.15, *p* = .03, *d* = .79. An additional independent-samples t-test with the average time that participants spent to solve the extinction trial as the dependent variable showed that participants in the experimental group took more time to complete the extinction trial (*M* = 107.22 seconds, *SD* = 94.33, *SE* = 22.23) than participants in the control group (*M* = 71 seconds, *SD* = 47.19, *SE* = 11.12). However, this comparison did not reach statistical significance, *t*(34) = 1.75, *p* = .09, *d* = .48. In line with the overall nature of these results, the verbal reports collected from the participants while working on the extinction trial revealed a sort of frustration experienced exclusively by participants in the experimental group: “This is impossible”; “I can’t understand how to solve it. This is the basis I need, but I am already very close to the budget”; “I can’t do anything”.

5.3 Role of mental set in moderating the subjective experience of the task

To assess the role of the mental set in moderating the subjective experience of the task, we compared how the two groups differed with regards to (a) their level of enjoyment of the task, (b) their perceived difficulty in the task, and (c) their awareness of how they had completed the task (i.e. in an automatic vs. reflexive way).

Overall, participants reported that they enjoyed playing the game (*M* = 4.11, *SD* = 0.14, *SE* = 0.14) and did not find the task difficult (*M* = 4.22, *SD* = 0.72, *SE* = 0.11).

Independent-samples t-tests also revealed that the experimental group and the control group did not differ in their level of enjoyment of the game (Experimental group: M = 4.17, SD = 0.86, SE = 0.20; Control group: M = 4.05, SD = 0.80, SE = 0.19), $t(34) = -.40$, p = .69, d = .14, and their perception of how difficult the task was (Experimental group: M = 4.17, SD = 0.62, SE = 0.14; Control group: M = 4.28, SD = 0.83, SE = 0.19), $t(34) = .45$, p = .65, d = .15.

Finally, a Chi-Square test with the proportion of participants who stated they had designed the bridges in an automatic way as the dependent variable revealed a significant difference between the two groups, $\chi^2(1) = 7.34$, $p = .007$ $\varphi = .46$, with more participants in the experimental group designing the bridges in an automatic way (9 out of 18, or 50%) compared to the control group (2 out of 18, or 11%). In line with this result, when at the end of the study participants were asked to clarify the type of strategy they adopted to design the bridges, those who had demonstrated mental set during the task often made statements such as "The only thing I thought was to construct a bridge that worked", or "I did not pay so much attention". In contrast, the participants who had *not* demonstrated mental set during the task often made statements such as "I had the feeling that you were trying to trick me", or "This is tricky because I don't have to use all the anchor points".

To assess participants' awareness of the research hypothesis, at the end of the experiment, we asked participants to guess the aim of the study. Most of them (23 out of 36, or 64%) were unable to guess the general class of phenomena that we were investigating (fixation, bias, mental blocks, etc.). Interestingly, the sub-group that was best able to guess the aim of the study mainly consisted of participants belonging to the experimental group (10 out of 13, or 77%) and they were those who had been the most stuck during the task (participants who guessed the aim of the study accounted for 65.72% of the measured ‘fixated’ designs).

6 DISCUSSION

With the objective of developing a mental set paradigm for studying design fixation, we developed a computer-based design task that would address some of the main limitations of previous design fixation studies, especially the exclusive focus on unbounded ideation, limited data capture and the use of subjective metrics to evaluate the design outcomes.

In line with previous psychological research targeting mental set [13, 14], we found that, when working on the dual-approach trials, and in comparison to the control group, those in the experimental group more often failed to notice the possibility of a less expensive design approach; they instead continued to design structures following the more expensive approach used in the previous single-approach trials. In other words, the development of a mechanized state of mind fixated the experimental group on a particular design approach, thus preventing the consideration of alternative, lower cost solutions. In addition, participants in the experimental group more often stated that they had constructed the bridges in an automatic way compared to the control group. Additional support for this claim came from their description of the strategy adopted to design the bridges (see Section 5.3). Similar statements were reported in Luchins’ original study (e.g. participants exhibiting the Einstellung effect said “I did not think of other methods because the same methods always worked”; “it became natural, automatic”;

participants not exhibiting the Einstellung effect said “I am not a fool”; “I am not that dumb”; “Are you trying to catch me?”) [14 (pp. 31, 32, 54)]. Overall, these observations support the claim that mental set occurs as an automatic repetition of familiar behavior [14, 15, 44].

Importantly, and contrary to previous design fixation studies, the fixating solution in our study was one which the participants arrived at spontaneously, rather than one which was directly provided by researchers. The provision of existing design solutions (e.g. in the form of an example solution) has often been shown to reduce the quantity and variety of the solutions that participants generate in response to a design task [11]. These experimental findings relate to observations of professional practice where knowledge of previous design solutions can ‘fixate’ the subsequent design behavior. For example, Eckert, Stacey, and Earl [35] state that when looking for solutions to specific problems, designers typically take inspiration from past designs. One consequence of this is that the identification and adaptation of prior work imports more into the new design than just a solution principle. It also carries assumptions about physical properties, materials, manufacturing processes and context of use. Whilst some of these assumptions may be required, others may be inappropriate, having been unintended in the new context and going unrecognized as the project progresses (for similar observations see [36, 37, 38]).

Just as external forms of inspiration can be a source of fixating knowledge, they are not the only source; prior experience in a particular task domain can also impact subsequent design behavior. Consistent with this, in our study, fixation (in the form of mental set) resulted from the participants’ experience with a particular design approach.

Similar to previous design fixation studies, our results are supported by observations of design practice in which the solution concepts developed in the early stages of the design process can have a limiting effect on later ideation, as effort is expended on defending the early direction rather than exploring new ones [5, 39, 40]. Ball and Evans [41] regarded such behavior as indicating a fixation on initial concepts, and a reliance on a simple ‘satisficing’ design strategy in contrast to any more ‘well-motivated’ process of optimization. In relation to this, many researchers have reported evidence for the existence of an ‘opportunistic’ behavior leading designers to base their decisions on familiar aspects of the task rather than on a hierarchically-structured top-down approach [42, 43, 44].

In light of these considerations, we believe that our study complements and enlarges previous research into the effect of external inspiration sources on the design outcome by providing new insights into the role of an ‘internal’ source of fixation taking the form of a premature (and unconscious) commitment to previously explored solutions or previously implemented approaches.

While examining design outcomes allowed us to evaluate the role of mental set in moderating the design solution, looking at the process of designing gave us the possibility to make observations about the nature of fixation episodes and the conditions in which they occurred. This was possible thanks to the use of a digital platform which captured design behavior *throughout* the task. We found that the occurrence of mental set in the experimental group led those participants to include more structural elements in their designs, and to spend more time designing the bridges. More interestingly, we observed that participants in the two groups showed different behaviors in their attempts to design the least expensive bridge that would support the load: the experimental group more often demonstrated a ‘strong then cheap’ behavior whilst the control group tended to demonstrate a ‘cheap then strong’ behavior. In light of this observation, we hypothesize a different role for ‘functional objectives’ (i.e. structural support) and ‘resource constraints’ (i.e. available budget) in moderating the design behavior of the two groups. While the behavior of the experimental group was mainly driven by the objective of designing a functional structure, the control group was more focused on the constrained resources, designing the least expensive bridge possible. The occurrence of these different patterns of behavior may have been induced by the different order of trials for the two groups, leading participants in the control group to approach the post-extinction trials by adopting a step-by-step approach (i.e. a step-wise increment of the bridges’ strength), ultimately resulting in more efficient designs. Another possibility is that the lower budget available to solve the extinction trial (see Table 1) might have altered participants’ design behavior in the following trials, to the extent that they became 'mentally set' to look for very inexpensive solutions. However, participants in both groups were instructed to design the least expensive bridge that could support the load and were motivated to do so by the prize available.

Our observations of the participants’ design process can be connected to Fricke’s accounts of early-stage design strategies [45]. He distinguished between a “function-oriented” strategy, in which the design operations are carried out for one initial function until a satisfying level of concretization is reached, and a step-wise “process-oriented” strategy, which follows a hierarchical and sequential plan of action, executing basic design operations step by step. In his study, Fricke noticed that, compared to the stepwise “process-oriented” strategy, the “function-oriented” strategy resulted in the generation of fewer solution variants. When combined with Fricke’s observations, our findings suggest that the strategy designers use while they are working on a problem impacts the type and variety of solutions they generate. This highlights a need for design fixation research to investigate (a) the influence of different design strategies on the occurrence of fixation episodes, and conversely (b) the influence of fixation episodes in determining the adoption of different design strategies. This may help researchers to develop a better understanding of the cognitive biases that can occur during the design process and allow recommendations to be formulated for which design strategies should be encouraged or discouraged, depending on either the nature of the design activity or the stages of the design process (see [18]).

In addition to contributing to the design literature, our findings expand previous results related to mental set in psychological research. At the most general level, the present study documented for the first time the occurrence of mental set by using a task that did not have a single best [14] or most efficient solution [13], but a range of possible more or less efficient responses. Participants were able to demonstrate flexibility both in how they approached the task and also in the solutions that they proposed. Indeed, an important difference between psychological research on fixation-related effects (in general) and design fixation studies (in particular) is that the former adopts highly constrained and well-defined tasks in which the solution has to be discovered, while the latter uses ill-defined and open-ended problems in which the solutions have to be invented [46, 47, 48]. Our task lies on this continuum: it was relatively constrained (i.e. to get an adequate control on the type of solutions participants could generate) and yet open-ended (to permit a design space to be explored). As such, our method and results might be of interest not only to those studying design fixation (who might see the opportunity to observe the phenomenon of interest in more constrained tasks) but also for those studying mental set (who might see the opportunity to study the phenomenon of interest in less constrained tasks).

In contrast to previous studies [13, 14], the occurrence of mental set did not prevent our experimental group from solving the extinction trial and only slightly slowed down their performance on that trial (for similar results on the extinction trial see [14]). One possibility could be that by allowing many solution variants we encouraged participants to expand the boundaries of the solution space, thus considering a larger set of solution approaches. Nevertheless, the verbal reports collected during the task revealed a sort of tension and nervousness experienced *exclusively* by participants in the experimental group when working on the extinction trial (see Section 5.2). These comments resembled those reported in Luchins’ original manuscript (e.g. “The rule doesn't work here,” “You made a mistake in this problem,” “The answer is 42 not 25”) [14 (p. 68)] and suggest that, even when the occurrence of mental set does not hinder the solution of the extinction trial, it may result in a different subjective experience of that trial. Furthermore, the fact that all the participants in the experimental group were able to complete the extinction trial confirmed that that group’s failure to use the alternative, less expensive design approach in the dual-approach trials occurred because of the blocking influence of the familiar approach and did not result from a lack of knowledge of how to apply that less expensive approach (also see [13]).

Looking at more specific results, we showed the ‘costs’ of mental set could be quantified in a different way to that demonstrated by Bilalić et al. [13]. In their study, participants completed only one 2-solutions trial and one 1-solution trial; the occurrence of mental set was quantified in terms of standard deviations in skill level on the extinction trial (i.e. the occurrence of mental set reduced the performance of chess players by three standard deviations to the level of less-skilled players). In contrast, following Luchins’ example [14], our participants completed four dual-approach trials and the effect of mental set was quantified in terms of the number of trials completed with the less efficient (i.e. more expensive) approach by participants in the two groups. From this perspective, our results expand those of Bilalić et al. as they give the opportunity to quantify mental set by evaluating the persistence of the effect across multiple trials and in tasks that permit multiple solutions.

In summary, we believe that our method promises a number of advantages over traditional design fixation studies:

- observation of design fixation beyond the context of idea generation;
- more objective analysis of design behavior (because the design task is partially constrained and outputs can be tested automatically);

- clear evaluation of the effects of objectives and constraints on design performance (because the objectives and constraints are well defined and performance is directly measured);
- direct quantification of the consequences of design fixation (because the inclusion of a resource objective allows researchers to measure the relative merits of different solutions);
- evaluation of design fixation episodes occurring during the design process (because using a digital platform allows researchers to conveniently capture and record design activities, not just the outputs);
- closer connection to theories, methods and results from psychological research literature (because the task and format more closely resemble work in that discipline).

In addition, providing feedback on design performance (through the possibility to switch between the Design and the Test mode) allowed us to simulate an important part of design activities that allow for rapid testing or simulation. This suggests a further advantage that our method may offer over the tasks typically used to study design fixation: the possibility for participants to attain a higher level of enjoyment and involvement in the task. This could easily be heightened by implementing a more explicit game format as previous research suggests that using a game-based approach can help develop and sustain engagement, compliance and satisfaction in the task (for a review see [49]). In addition, playfulness and task motivation are known to be important factors in fostering creativity [50, 51]. From this perspective, using computer games that represent simple design activities may allow fixation researchers to either achieve a greater level of participant involvement in the task (thus avoiding the limitations frequently highlighted in design fixation studies, e.g. [14, 52, 53]) or increase the creative potential of the participants’ responses.

Finally, our observation that participants who experienced fixation during the design activity increased their awareness of the occurrence of this phenomena highlights some potential for using a variant of our method to encourage designers to recognize fixation effects and possibly overcome them (for a review of research on computer-based game-like tasks as a means to promote behavioral and attitudinal changes see [54]). Support for this possibility comes from research both within the design field [47] and elsewhere [55, 56, 57], suggesting that experience of fixation itself (and its negative consequences) can be the means by which designers reflect on their biases and learn to resist them.

7 LIMITATIONS AND FUTURE WORK

Despite the benefits of our approach to studying design fixation, there are some limitations and challenges that require attention. In particular, when selecting or developing suitable tasks, it is difficult to arrive at tasks that are adequately controlled and yet flexible enough to permit creative design work to be performed. As mentioned before, in developing our study we tried to circumvent this limitation by using a task which was quite constrained (i.e. to gain an adequate control on the performance) but permitted a range of possible responses (to permit a design space to be explored). Nevertheless, we are aware that the use of relatively constrained computer-based tasks in fixation research calls for balancing several factors, including the complexity of the design tasks, the number of possible solutions to those tasks and the time taken for participants to learn how to use the software. Future work might progressively vary the degree to which the task is constrained (and thus the range of possible solutions) and assess how this influences the number and type of solutions participants generate.

Our task exhibited some very specific features (i.e. type of structure that could be constructed, type of materials that could be used) and the type of design activity required. This raises a number of questions about how the findings from our research might generalize to other kinds of design activity, and also questions about what kind of design tasks offer the best basis for studying fixation effects at different stages of the design process. Furthermore, our design task had a relatively simple set of objectives and constraints (the bridge should span the river, be sufficiently strong and yet be as cheap as possible). Designers often work with numerous interacting and conflicting requirements, not all of which are well defined or directly reducible to numerical values. Future studies could take inspiration from our approach to develop novel tasks that can be used to study how fixation episodes occur in various types of design activities and at the various stages of the design process. In developing such tasks, researchers should be aware that mental set may not always be easily induced experimentally, at least not in the strongest form that Luchins achieved. In particular, we recommend that the 'fixating' trials be selected carefully (possibly through a pre-testing phase) so that all the participants are able to solve them. Indeed, if these trials are too hard, participants may not be able to complete them and thus mental set may be not induced [58].

Design fixation studies typically permit participants to sketch (freehand) in response to the task that is set, potentially imitating the practices of some working designers [59 (p. 25)], or encouraging creative behavior [60]. In contrast, our study required participants to work directly on the computer, allowing the provision of a design environment in which participants received feedback on their design performance (through the possibility to switch between the Design and the Test mode). This mimics an important part of design activities that allow for rapid testing (e.g. design practices in which computer simulations are common: Finite Element Analysis, Computational Fluid Dynamics, etc.). However, whilst many practicing designers work only (or primarily) with digital tools [61], many do not, especially in early-stage design. Despite the methodological benefits of constraining participants to the use of computers, this might affect the results when certain research questions are being investigated (e.g. those connected to modes of representation). Future work might then develop a hybrid approach, where participants are able (or required) to sketch by hand before entering a digital environment that permits feedback and iteration.

Despite the differences between our study and a conventional design fixation experiment, in some ways it is also quite similar, especially with respect to the characteristics of the participants (inexperienced, unspecialized), the duration of the task (short) and the discipline of design (Engineering). Whilst these features might all pose problems for generalizing our results to other kinds of design practice, that is not our objective here. We only seek to make claims about methodological options, and the type of method we demonstrate is equally applicable to expert designers working on longduration tasks. Furthermore, the opportunity to administer the study remotely (e.g. online) and thus periodically (e.g. once a week) would make such studies easier to conduct. Of course, our particular task only relates to structural design, but future research might explore mental set in different types of design activities. All those design practices that involve repeatedly solving similar but different design problems might be subject to mental set, including design work that produces plans for similar but different structures, mechanisms, electrical circuits and software routines. In such cases, it could be that an implicit assumption leads designers to repeat a category of solution or the means by which that solution is reached. Investigating mental set in engineering practice might either reveal differences in how fixation is manifested in different problem types, or might reveal that the phenomena of interest are in fact quite similar. As such, researchers could select their experimental tasks based on the methodological opportunities those tasks offer rather than on some similarity to specific design practices.

8 CONCLUSION

Over 25 years, design fixation has provided researchers from a variety of backgrounds with a compelling, important, and uniquely cross-disciplinary design phenomenon to study. The research is compelling because fixation blocks the creativity of designers, thus limiting their ability to generate innovative solutions. Individually and collectively, studies of design fixation have improved our understanding of why fixation occurs and how it might be mitigated. However, to date, such studies have been quite uniform in the experimental approach adopted. That approach suffers from a number of methodological limitations related to the way in which the occurrence of fixation episodes is identified (i.e. at early-stage ideation), the characteristics of the design problems (i.e. ill-defined and ill-structured problems), the type of data that are collected (i.e. participants’ final design ideas) and the way in which the data are analyzed (i.e. through various subjective metrics). The method we propose here provides a promising future direction for fixation research, offering a more objective, repeatable and comparable description of the various phenomena of interest. In conclusion, although design fixation research has already made good progress with a very limited set of experimental techniques, there are great opportunities for developing other approaches. Applying a broader range of experimental methods might be expected to generate richer insights into fixation, how it occurs and how it might be mitigated.

ACKNOWLEDGMENT

For their feedback on this work, we would like to thank Prof. Merim Bilalić (Department of Psychology, Northumbria University), Dr Sanne de Wit and Dr Aukje Verhoeven (Department of Clinical Psychology, University of Amsterdam), and Dr Carlos Cardoso and members of the Product Innovation Management department (Industrial Design Engineering, Delft University of Technology). We are also grateful to three anonymous reviewers for their constructive comments on an earlier version of this paper.

FUNDING

This work was supported by the UK's Engineering and Physical Sciences Research Council (EP/K008196/1) and the CAPES Foundation Ministry of Education of Brazil (BEX 11468/13-

0). Research data supporting this publication are available from the University of Cambridge data repository at https://doi.org/10.17863/CAM.7957.

REFERENCES

[1] Youmans, R. J., & Arciszewski, T. (2014). “Design fixation: classifications and modern methods of prevention”. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. **28**(02), pp. 129-137.

[2] Youmans, R. J., & Arciszewski, T. (2014). “Design fixation: a cloak of many colors”. In Design Computing and Cognition’12 (pp. 115-129). Springer: The Netherlands.

[3] Perttula, M., & Sipilä, P. (2007). “The idea exposure paradigm in design idea generation”. Journal of Engineering Design, **18**(1), pp. 93-102.

[4] Agogue, M., Poirel, N., Pineau, A., Houd e, O., & Cassotti, M. (2014). “The impact of age and training on creativity: a design-theory approach to study fixation effects”. Thinking Skills and Creativity, **11**, pp. 33-41.

[5] Crilly, N. (2015). “Fixation and creativity in concept development: The attitudes and practices of expert designers”. Design Studies, **38**, pp. 54-91.

[6] Robertson, B. F., & Radcliffe, D. F. (2009). “Impact of CAD tools on creative problem solving in engineering design”. Computer-Aided Design, **41**(3), pp. 136-146.

[7] Jansson, D. G., & Smith, S. M. (1991). “Design fixation”. Design Studies, **12**(1), pp. 3-11.

[8] Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L., & Schunn, C. (2010). “A study of design fixation, its mitigation and perception in engineering design faculty”. Journal of Mechanical Design, **132**(4), 041003.

[9] Viswanathan, V., & Linsey, J. (2011). “Design fixation in physical modeling: an investigation on the role of sunk cost”. In ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Washington, USA, pp. 119-130.

[10] Vasconcelos, L. A., & Crilly, N. (2016). “Inspiration and fixation: Questions, methods, findings, and challenges”. Design Studies, **42**, pp. 1-32.

[11] Sio, U. N., Kotovsky, K., & Cagan, J. (2015). “Fixation or inspiration? A meta-analytic review of the role of examples on design processes”. Design Studies, **39**, pp. 70-99.

[12] Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). “Toward a conceptual framework for mixed-method evaluation designs”. Educational evaluation and policy analysis, **11**(3), pp. 255-274.

[13] Bilalić, M., McLeod, P., & Gobet, F. (2008). “Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect”. Cognition, **108**(3), pp. 652-661.

[14] Luchins, A. S. (1942). “Mechanization in problem solving-the effect of Einstellung”. Psychological Monographs, **54**, No. 6: 1-95.

[15] Simon, H. A., & Gilmartin, K. (1973). “A simulation of memory for chess positions”. Cognitive Psychology, **5**(1), pp. 29-46.

[16] La Rocca, G., & van Tooren, M. J. (2010). "Knowledge-based engineering to support aircraft multidisciplinary design and optimization". Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, **224**(9), pp. 1041-1055.
[17] Snider, C. M., Culley, S. J., & Dekoninck, E. A. (2013). "Analysing creative behaviour in the later stage design process". Design Studies, **34**(5), pp. 543-574.
[18] Snider, C., Dekoninck, E., & Culley, S. (2016). "Beyond the concept: characterisations of later-stage creative behaviour in design". Research in Engineering Design, **27**(3), pp. 1-25.
[19] Maier, N. R. (1931). "Reasoning and learning". Psychological review, **38**(4), 332.
[20] Duncker, K., & Lees, L. S. (1945). "On problem-solving". Psychological monographs, **58**(5), pp. 1-113.
[21] Luchins, A. S. (1951). "On recent usage of the Einstellung-effect as a test of rigidity". Journal of Consulting Psychology, **15**(2), pp. 89-94.
[22] Luchins, A. S., & Luchins, E. H. (1950). "New experimental attempts at preventing mechanization in problem solving". The Journal of General Psychology, **42**(2), pp. 279-297.
[23] Luchins, A. S., & Luchins, E. H. (1959). "Rigidity of behavior: A variational approach to the effect of Einstellung". University of Oregon Books
[24] Chen, Z., & Mo, L. (2004). "Schema Induction in Problem Solving: A Multidimensional Analysis". Journal of Experimental Psychology: Learning, Memory, and Cognition, **30**, pp. 583-600.
[25] Ellis, J. J., & Reingold, E. M. (2014). "The Einstellung effect in anagram problem solving: evidence from eye movements". Frontiers in Psychology, **5**(679), pp. 1-7.
[26] Lovett, M. C., & Anderson, J. R. (1996). "History of success and current context in problem solving: Combined influences on operator selection". Cognitive psychology, **31**(2), pp. 168-217.
[27] Atwood, M. E., & Polson, P. G. (1976). "A process model for water jug problems". Cognitive Psychology, **8**(2), pp. 191-216.
[28] Delaney, P. F., Ericsson, K. A., & Knowles, M. E. (2004). "Immediate and Sustained Effects of Planning in a Problem-Solving Task". Journal of Experimental Psychology: Learning, Memory, and Cognition, **30**, pp. 1219-1234.
[29] Tresselt, M. E., & Leeds, D. S. (1953). "The Einstellung effect in immediate and delayed problem-solving". The Journal of General Psychology, **49**(1), pp. 87-95.
[30] Bilalić, M., McLeod, P., & Gobet, F. (2010). "The mechanism of the Einstellung (set) effect: A pervasive source of cognitive bias". Current Directions in Psychological Science, **19**(2), pp. 111-115.
[31] Bilalić, M., & McLeod, P. (2014). "Why good thoughts block better ones". Scientific American, **310**(3), pp. 74-79.
[32] Saariluoma, P. (1992). "Error in chess: The apperception-restructuring view". Psychological Research, **54**, pp. 17-26.
[33] Chronic Logic (2001). Pontifex. http://www.chroniclogic.com/pontifex.htm. Accessed date: 21/04/2016.
[34] Beepa (2013). Fraps. http://www.fraps.com/. Accessed date: 21/04/2016.

[35] Eckert, C. M., Stacey, M., & Earl, C. (2005). "References to past designs". In J. S. Gero, & N. Bonnardel (Eds.), "Studying designers" (pp. 3-21). Aix-en-Provence, France: University of Provence.
[36] Busby, J. A., & Lloyd, P. A. (1999a). "Influences on solution search processes in design organisations". Research in Engineering Design, **11**(3), pp. 158-171.
[37] French, M. J. (1998). "Conceptual design for engineers" (3rd ed.). London, UK: Springer. Originally published in 1971 as "Engineering Design: The Conceptual Stage".
[38] Frensch, P. A., & Sternberg, R. J. (1989). "Expertise and intelligent thinking: when is it worse to know better?" In R. J. Sternberg, (Ed.) "Advances in the psychology of human intelligence", Vol. 5 (pp. 157-188). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
[38] Rowe, P. G. (1987). "Design thinking". Cambridge, MA: MIT Press.
[40] Ullman, D. G., Dietterich, T. G., & Stauffer, L. A. (1988). "A model of the mechanical design process based on empirical data". Artificial Intelligence for Engineering, Design, Analysis and Manufacturing, **2**(01), pp. 33-52.
[41] Ball, L. J., Evans, J. S. B., & Dennis, I. (1994). "Cognitive processes in engineering design: A longitudinal study". Ergonomics, **37**(11), pp. 1753-1786.
[42] Bender, B., & Blessing, L. (2004). "On the superiority of opportunistic design strategies during early embodiment design". In DS 32: Proceedings of DESIGN 2004, the 8th International Design Conference, Dubrovnik, Croatia.
[43] Cross, N. (2004). "Expertise in design: an overview". Design Studies, **25**(5), pp. 427-441.
[44] Guindon, R. (1990). "Designing the design process: Exploiting opportunistic thoughts". Human-Computer Interaction, **5**(2), pp. 305-344.
[45] Fricke, G. (1999). "Successful approaches in dealing with differently precise design problems". Design Studies, **20**(5), pp. 417-429.
[46] Bronowski, J. (1985). "The creative process". Leonardo, **18**(4), pp. 245-248.
[47] Crilly, N. (2010). "The roles that artefacts play: technical, social and aesthetic functions". Design Studies, **31**(4), pp. 311-344.
[48] Kuhn, T. S. (1962). "The structure of scientific revolutions". Chicago, Illinois: University of Chicago press.
[49] Lankoski, P., & Björk, S. (2015). "Game research methods: An overview". ETC Press, Carnegie Mellon University, Pittsburgh, PA.
[50] Amabile, T. M. (1983). "The social psychology of creativity: A componential conceptualization". Journal of personality and social psychology, **45**(2), p. 357.
[51] Howard-Jones, P., Taylor, J., & Sutton, L. (2002). "The effect of play on the creativity of young children during subsequent activity". Early Child Development and Care, **172**(4), pp. 323-328.
[52] Youmans, R. J. (2011). "The effects of physical prototyping and group work on the reduction of design fixation". Design Studies, **32**(2), pp. 115-138.
[53] Welch, M., Barlex, D., & Lim, H. S. (2000). "Sketching: Friend or foe to the novice designer?" International Journal of Technology and Design Education, **10**(2), pp. 125-148.

[54] Michael, D., & Chen, S. (2006) “Serious games: Games that educate, train, and inform”. Boston, MA: Thomson Course Technology.
[55] Croskerry, P. (2003). “The importance of cognitive errors in diagnosis and strategies to minimize them”. Academic Medicine, **78**(8), pp. 775-780.
[56] Kassirer, J. P., & Kopelman, R. I. (1989). “Cognitive errors in diagnosis: instantiation, classification, and consequences”. The American Journal of Medicine, **86**(4), pp. 433-441.
[57] Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). “Debiasing decisions: Improved decision making with a single training intervention”. Policy Insights from the Behavioral and Brain Sciences, **2**(1), pp. 129-140.
[58] Beilock, S. L., & DeCaro, M. S. (2007). “From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure”. Journal of Experimental Psychology: Learning, Memory, and Cognition, **33**(6), pp. 983-998.
[59] Henderson, K. (1999). “On line and on paper: Visual representations, visual culture, and computer graphics in design engineering”. Cambridge, MA: MIT press.
[60] Goldschmidt, G. (2014). “Modeling the role of sketching in design idea generation”. In A. Chakrabarti & L. T. M. Blessing (Eds.) “An Anthology of Theories and Models of Design” (pp. 433-450). London, UK: Springer.
[61] Chandrasegaran, S. K., Ramani, K., Sriram, R. D., Horváth, I., Bernard, A., Harik, R. F., & Gao, W. (2013). “The evolution, challenges, and future of knowledge representation in product design systems”. Computer-Aided Design, **45**(2), pp. 204-228.

Figures Captions List

Tables Captions List

Table 1 Characteristics of the trials and the trial order in the two groups. The column labelled "number of anchor points provided/minimum/maximum" details the number of anchor points provided to participants at the start of the trial, the minimum number of anchor points that must be used to span the gap and support the load, and the maximum number of anchor points that can be used to span the gap and support the load without exceeding the allocated budget. (For single-approach trials, provided = minimum = maximum; for dual-approach trials, provided = maximum > minimum; for the extinction trial, minimum = maximum < provided.)

Table 2 Number of participants using all the anchor points in the dual-approach trials and number of dual-approach trials solved using all the anchor points (percentages in brackets).

Table 3 The time spent designing across all the trials, mean number of structural elements used and number of design-test cycles (Standard Deviation and Standard Error in brackets).

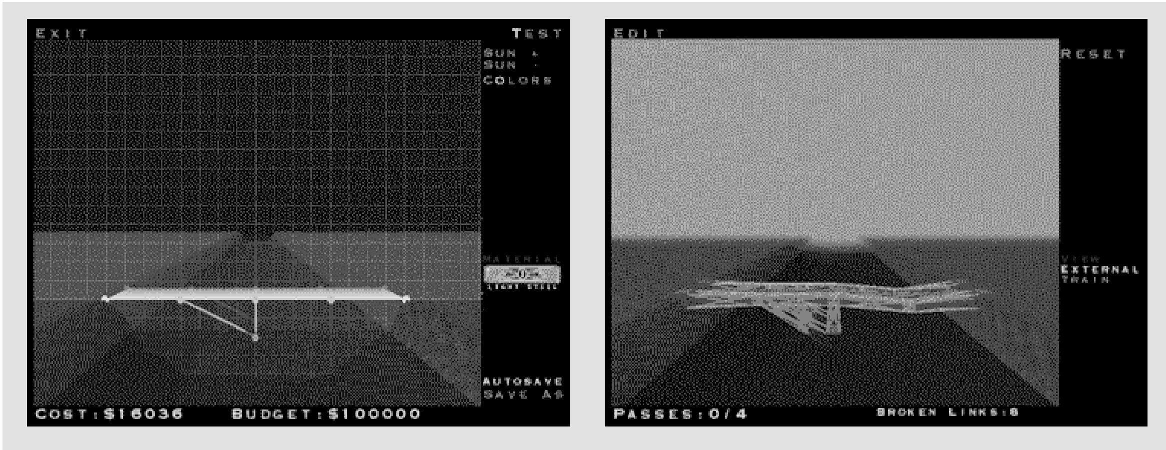

Figure 1. Screen images of the game platform used for the trials, showing the Design mode (left) and the Test mode (right).

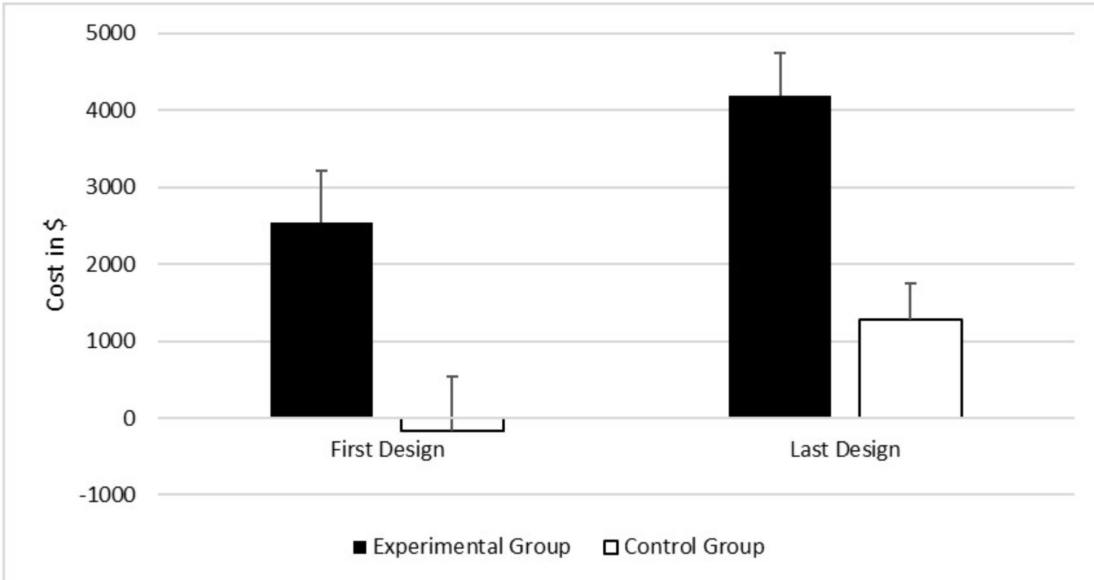

Figure 2. Mean additional cost (and Standard Error) of the last design tested compared to the first design tested. In each case, the additional cost is relative to the cost of the least expensive possible functional solution.

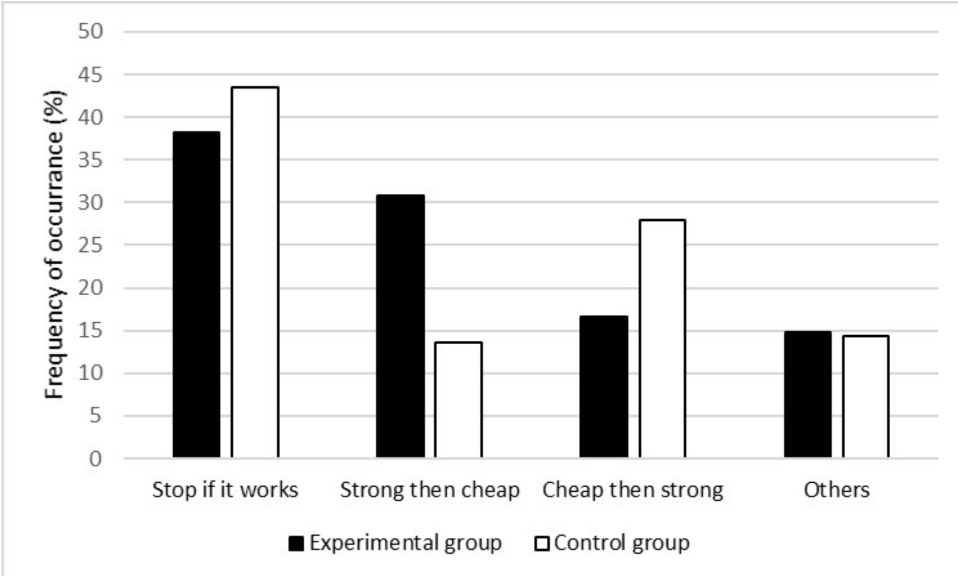

Figure 3. Proportion of trials in which the participants exhibited the ‘stop if it works’ behavior, the ‘strong then cheap’ behavior and the ‘cheap then strong’ behavior. ‘Others’ refer to all the cases in which the behavior could not be classified in one of previous categories or resulted from a combination of them.

Table 1. Characteristics of the trials and the trial order in the two groups. The column labelled "number of anchor points provided/minimum/maximum" details the number of anchor points provided to participants at the start of the trial, the minimum number of anchor points that must be used to span the gap and support the load, and the maximum number of anchor points that can be used to span the gap and support the load without exceeding the allocated budget. (For single-approach trials, provided = minimum = maximum; for dual-approach trials, provided = maximum > minimum; for the extinction trial, minimum = maximum < provided.)

Trial type	**Required bridge span**	**Budget available**	**Trials order in *experimental condition***	**Trials order in *control condition***	**Number of anchor points provided/ minimum/ maximum**	**Cost of the least expensive solution if…**	
						…using all the anchor points	**…not using all the anchor points**
Single-approach	160m	$100,000	1	3	3/3/3	$18,272	Bridge fails
Single-approach	240m	$100,000	2	9	5/5/5	$26,672	Bridge fails
Single-approach	240m	$100,000	3	10	5/5/5	$27,672	Bridge fails
Single-approach	320m	$100,000	4	6	6/6/6	$37,478	Bridge fails
Single-approach	320m	$100,000	5	4	5/5/5	$37,544	Bridge fails
Dual-approach	160m	$100,000	6	2	4/3/4	$18,772	$18,272
Dual-approach	320m	$100,000	7	5	7/6/7	$40,044	$37,544
Extinction	80m	$7000	8	1	4/2/2	Budget exceeded	$6400
Dual-approach	240m	$100,000	9	8	6/5/6	$27,672	$26,672
Dual-approach	240m	$100,000	10	7	8/6/8	$28,606	$26,606

Table 2. Number of participants using all the anchor points in the dual-approach trials and number of dual-approach trials solved using all the anchor points (percentages in brackets).

	Number of *participants* solving the dual-approach trials using all the anchor points	**Number of dual-approach *trials* solved using all the anchor points**
Experimental Group	15 (83%)	34 (47%)
Control Group	8 (44%)	13 (18%)

Table 3. The time spent designing across all the trials, mean number of structural elements used and number of design-test cycles (Standard Deviation and Standard Error in brackets).

	Mean time (in secs) spent designing	**Mean number of structural elements used**	**Mean number of design-test cycles**
Experimental Group	102.46 (*SD* = 43.41, *SE* = 10.23)	26.49 (*SD* = 8.79, *SE* = 2.07)	2.44 (*SD* = 0.83, *SE* = 0.20)
Control Group	74.44 (*SD* = 36.19, *SE* = 8.53)	19.61 (*SD* = 7.21, SE = 1.70)	2.23 (*SE* = 0.19, *SD* = 0.83)