The effect of explicit instructions in idea generation studies

Luis A. Vasconcelos, Maria A. Neroni, and Nathan Crilly

Engineering Design Centre, Department of Engineering, University of Cambridge

Corresponding author: Luis A. Vasconcelos

Mailing address: Engineering Design Centre, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, UK

Phone: 01223 748573

E-mail: lal40@cam.ac.uk

Short title: Instructions in ideation studies

Number of pages: 48

Number of tables: 10

Number of figures: 4

The effect of explicit instructions in idea generation studies

Abstract: In inspiration and fixation experiments, example designs are often provided along with instructions for how participants should treat them. However, research has not reached a consensus about the influence of such instructions, leading to difficulties in understanding how the examples and the instructions each affect idea generation. We conducted an experiment in which 303 participants designed for the same design problem, whilst given different examples and instructions, which ranged from strongly-encouraging copying the examples, to strongly-discouraging copying. Exposure to the examples affected the number and type of ideas generated, whereas exposure to the instructions did not. However, instructions did affect how participants incorporated features of the examples in their ideas. Encouraged groups incorporated many features of the examples, while also incorporating structural features more than conceptual ones. Surprisingly, the incorporation of features in discouraged groups was not different from that of groups given no instructions or even no stimulus. This indicates that concrete features may be easier to recognise and reproduce than abstract ones, and that encouraging instructions are more effective than discouraging ones, despite how strict or lenient those instructions are. The manipulation of different features also allowed us to observe how similar approaches to solve a design problem can compete for attention and how the calculation of feature repetition can be misleading depending on how common or obvious the features might be. These findings have implications for the interpretation of results from fixation studies, and for the development of design tools that present stimuli to assist idea generation.

Keywords: inspiration, design fixation, experimental instructions, feature repetition, stimuli introduction

1. INTRODUCTION

Inspiration is vital to creative design and innovation; yet, we know very little about it. Understanding more about inspiration and other cognitive aspects of design allows us to improve design processes and outcomes. It also allows us to develop intelligent tools to assist innovative design and computational models of creativity. The potential benefits of understanding inspiration have driven the design research community to conduct many studies that seek to find out, for instance, what materials inspire designers (Gonçalves et al., 2014), how designers achieve inspiration (Zhao, 2013), how inspiration can improve designers' performance (Eckert, 1997), and how to better develop computational tools to inspire designers (Töre Yargin & Crilly, 2015). Many of these studies have observed the use of external stimuli during idea generation, and have reported that whilst stimuli can assist idea generation, they can also hinder it. One of these hindrances to creativity is described in the design literature as "design fixation" (Jansson & Smith, 1991), when prior knowledge of a solution causes designers to inadvertently limit their search for other solutions.

Design fixation was originally studied in situations where designers were trying to address a problem and were provided with stimuli which represented possible solutions to the problem that was being addressed (Jansson & Smith, 1991). Fixation was measured by counting the solutions that were developed in response to the problem and counting the repetition of features from the stimuli. This general approach to measuring fixation has subsequently been employed in several other related studies (e.g. Purcell & Gero, 1996; Dahl & Moreau, 2002; Linsey et al., 2010; Cardoso & Badke-Schaub, 2011; Youmans, 2011a, 2011b, Vasconcelos et al., 2017). Although reproducing features of good designs might be beneficial or efficient, even flawed stimuli (i.e. stimuli with negative features) are reproduced, when it might instead be expected that those stimuli would be identified and avoided (Jansson & Smith, 1991; Linsey et al., 2010; Youmans, 2011b).

In general, fixation research suggests that blindly copying features from stimuli is harmful to idea generation, and therefore to the design process itself. As such, some studies have incorporated constraining textual instructions into the stimuli given to participants, to prevent them from copying their features. However, the efficacy of these instructions has varied across studies: they were effective in some cases (Chrysikou & Weisberg, 2005; Yilmaz et al., 2010) and ineffective in others (Jansson &

Smith, 1991; Perttula & Sipilä, 2007). Thus, it is still uncertain how textual instructions can influence the copying behaviour, and this uncertainty might be attributed to different factors. For instance, individuals may tend to overlook the instructions that accompany the stimuli (LeFevre & Dixon, 1986), they may interpret the instructions in different ways if instructions are not explicit, or the instructions may lead them to question why they are being exposed to the stimuli and change their idea generation process accordingly. Whatever the reason might be, it is important to understand the relevance of the instructions provided to designers as part of the inspiration material. Methodologically, this would help in determining how fixation studies should be conducted and the results interpreted. More practically, it would also help in understanding how stimuli should be presented when designers are stimulated with examples, such as in those software tools that support idea generation by providing external stimuli (for examples of such tools, see (Chakrabarti et al., 2005; Linsey & Wood, 2007; Vattam et al., 2011)).

To better understand the influence of instructions on idea generation, we conducted an experiment in which participants responded to the same design problem, whilst being exposed to different example solutions, and following different instructions with respect to those examples.

2. INSPIRATION, FIXATION, AND THE INTRODUCTION OF EXTERNAL STIMULI

Design researchers have studied many different aspects of the inspiration process, including characteristics of external stimuli or inspiration sources (e.g. the modality of representation used for the stimuli) (Linsey et al., 2008; Sarkar & Chakrabarti, 2008; Atilola & Linsey, 2015; Viswanathan et al., 2016) and aspects of the design process (e.g. time available for ideation) (Youmans, 2011a; Tsenn et al., 2014; Siangliulue et al., 2015). One of the stimuli characteristics that has been studied is whether the quality of the examples (i.e. good or bad) would influence feature repetition, or whether those examples would be copied indiscriminately. A series of studies has found that participants still repeat features from previous examples even when they are flawed (Jansson & Smith, 1991; Chrysikou & Weisberg, 2005; Perttula & Liikkanen, 2006; Fu et al., 2010; Linsey et al., 2010; Youmans, 2011b; Viswanathan et al., 2014). It should be noted that in Jansson and Smith's (1991) original paper, feature repetition was a way in which the researchers operationalised the degree to which the participants' exploration of the solution-space was limited. Furthermore, the repetition of *negative* features (or features that contradicted the problem brief) was taken to be an indication that this limited search was an unconscious behaviour (Youmans & Arciszewski, 2014).

To counteract the indiscriminate repetition of features, some studies have tried warning participants about the flaws in the examples (Jansson & Smith, 1991; Chrysikou & Weisberg, 2005; Viswanathan et al., 2014), while others have tried instructing participants not to copy the examples (Smith et al., 1993; Perttula & Sipilä, 2007; Yilmaz et al., 2010). Considering these two approaches, Chrysikou and Weisberg (2005) found that warning participants of flaws in the examples was not enough; they had to be told to avoid repeating those flaws. Yilmaz et al. (2010) also told participants not to reproduce the examples and this led to a reduction in feature repetition. Conversely, Jansson and Smith (1991) and Perttula and Sipilä (2007) found repetition even when participants were instructed to avoid using features of the examples provided. In summary, the published literature reports conflicting results about the influence of instructions when providing external stimuli to designers.

Whilst many variables have been manipulated in fixation experiments, and some studies have already tested the effectiveness of using textual instructions to some extent, the way the stimuli are introduced in such experiments has not yet been studied systematically. Such stimuli introductions can typically be divided into two components: a descriptive statement on what the stimulus is (e.g. "*here is an example solution*") and a prescriptive instruction for how the stimulus should be used (e.g. "*don't copy its features*"). Currently, the stimuli introductions (i.e. descriptions and instructions) given to participants vary from study to study, with some reporting that the example is there to help, to illustrate previous solutions, or to show how ideas should be represented, while other studies only mention that the stimulus was provided to participants without describing how (Vasconcelos & Crilly, 2016). See Table 1 for a more detailed description of stimuli introductions as described in related studies.

Table 1. Different means of introducing external stimuli to participants (each instance was extracted from a fixation-related study; changes to the original text were made to permit better comparison across rows)

Study	**Stimuli introduction (or descriptive statement about the stimuli)**
(Cheng et al., 2014)	The examples should be analysed using a set of instructions
(Dahl & Moreau, 2002)	The example is provided to help participants get started and it might be useful in solving the design problem
(Gonçalves et al., 2012)	The example could be considered (or not) when participants were generating ideas
(Liikkanen & Perttula, 2008)	The example should be used to raise thoughts
(Linsey et al., 2010)	The example should be considered as a solution that might be created for the design problem
(Lujun, 2011)	The example could be referred to during ideation
(Perttula & Sipilä, 2007)	The examples should be used to awaken thoughts, but should not be reproduced as such
(Purcell & Gero, 1996)	The example design illustrates what is meant by a sketch design
(Vasconcelos et al., 2016)	The example shows how participants should present their ideas
(Vasconcelos et al., 2017)	The example is a concept that illustrates one way to solve this problem
(Yilmaz et al., 2010)	The examples could be used to understand the method, but they should not be repeated in participants' own designs
(Youmans, 2011a)	The example illustrates a previous design for the same task
(Youmans, 2011b)	The example should be rebuilt as to allow familiarization with the materials in a construction set

Variation in the way the stimuli are introduced might be attributed to a lack of agreement across studies about which "real world" situation is being simulated (e.g. contexts in which examples are accidentally seen, intentionally searched for, or already known). Regardless of the reason, the variation in the way copying is encouraged or discouraged makes it difficult to compare results across studies and to know the extent to which the design work is being affected by instructions.

Although previous fixation research makes it hard to formulate a theory for the influence of instructions, there is one study that looked more carefully at how different sets of instructions affected feature repetition during idea generation (Smith et al., 1993). In the experiment, participants were asked to either conform to or diverge from the example solutions provided in a creative task. Some participants (conform group) were told that the examples were great ideas previously created for that task, and that participants should create ideas like those whilst not copying them exactly. Other participants (diverge group) were told that the examples restricted people's creativity, and that participants should create ideas that were different from the examples. Finally, other participants (standard group) were given the same examples but without any specific instructions, and a baseline group was given no examples at all. When compared to the participants in the baseline group, the researchers found that all other groups generated more ideas containing features of the examples. Additionally, feature repetition in the diverge group did not decrease when compared to the standard group, while feature repetition in the conform group increased when compared to the standard group.

Based on their results, the researchers proposed that the participants did not conform to examples because they had assumed that they should; they conformed because they could not forget the examples that they had seen. However, the instructions used in the experiment were not strict (i.e. they suggested how ideas should be created, but neither forbade nor required the use of features from the examples, thus allowing participants to interpret the instructions in different ways). The description of the stimuli and the instruction for their use also varied between groups, making it difficult to infer the influence of each piece of information in isolation. Additionally, the experiment reports an analysis on the repetition of features that were always concrete or structural (e.g. parts of a physical object), thus inviting further investigation to be done on different types of features, such as more abstract or conceptual features (e.g. properties of a physical object).

We believe that a different experimental setup with additional manipulations can provide further evidence to clarify the influence of the textual instructions used in experimental research, thus having implications for design research methodology and for professional design practice. To investigate this, we designed and conducted a study to explore the effects of instructions in design inspiration and fixation experiments, with instructions that are (slightly or very) discouraging or encouraging, or that are neutral.

3. RESEARCH METHODOLOGY

We conducted a single experiment ($N = 303$) designed to examine the role of instructions in creativity studies. However, whilst the data was captured in a single session (thus minimising experimental variation and allowing for efficient data capture), the experimental manipulation and metrics were divided into two main blocks, and the corresponding analyses were intended to be performed over two different time frames. Thus, we here treat all the data as relating to two separate studies; the details and results of which are discussed as follows. The experimental setup of Study 1 is described in full, whereas Study 2 follows up on that with shorter descriptions, in which only the main differences from the previous study are highlighted. Both studies drew participants from the same population (e.g. educational background), and used identical procedures (e.g. time available), materials (e.g. design problem), analysis (e.g. fixation metrics), and environment (e.g. room conditions). However, the stimulus manipulated in Study 2 incorporated a distinct conceptual feature, thus being significantly different from that used in Study 1.

3.1 Objective and hypothesis

This experiment investigates how textual instructions accompanying external stimuli can shape the design work of the participants. We hypothesise that instructions for the incorporation of features provided along with an example design will determine the repetition in the ideas that participants generate. In particular, we expect that positive or encouraging instructions will increase feature repetition, whereas negative or discouraging instructions will reduce feature repetition. Additionally, we also expect to find some level of fixation in all groups exposed to the examples when compared to a baseline group. From an experimental perspective, we understand design fixation to be the process of generating ideas in response to a design problem under the influence of one or more external stimuli, in which the generation of ideas is constrained due to the repetition of the stimuli into the ideas generated. This would imply decreased productivity of the participants as well as increased idea repetition.

4. STUDY 1

4.1 Experimental method

Participants were randomly allocated to different experimental conditions. They were verbally asked to be creative and to generate, individually, as many ideas as possible to a given problem. They were also instructed, via voice and text, to sketch and describe in writing their ideas on sheets of paper. Except for a baseline group that designed without any stimuli, all other experimental groups received a sketch of one example solution and a description of what the sketch represented. These stimulated groups received instructions for the use of features from the example, and these instructions varied with respect to how constraining or encouraging they were (see the materials section for the complete instructions). Finally, the participants' ideas were assessed to evaluate the influence of the instructions on the level of fixation observed.

4.2 Participants

One hundred and sixty-eight engineering students in their first year at the University of Cambridge (UK) were assigned to six experimental groups (n = 28). Participation in the experiment was part of the students' education (but was not for credit), and was aimed at collecting data that could later be used to introduce them to the concept of design fixation. No demographic data was collected from the participants, but as first year undergraduate students they were broadly similar in age and design experience, drawn from a cohort with a male-female ratio of approximately 3:1.

4.3 Task and problem

The participants were told to solve the following problem. "*Bicycles are a popular mode of transportation and recreation for many people. While growing up, a person might go through a series of ever-larger bikes, sometimes having several models, one after the other. However, having several bikes can be a problem for many reasons. Your task is to generate as many ideas as possible to eliminate the need to have multiple bikes as people grow up.*"

This problem was selected because it was expected to satisfy four criteria. First, it was unlikely that the participants had designed solutions for this problem before. Second, they were likely to have experienced the situation described in the problem previously (i.e. while growing up, they probably had multiple bikes), thus helping their understanding of it. Third, the problem could be solved in different ways, with several different underlying principles being applied, thus leaving enough room for creativity. Finally, the design brief held a low level of complexity, which was also expected for the ideas generated, thus being suitable for a quick experiment fitting with the constraints of the session.

4.4 Procedure (overview)

The experiment took place in a large lecture theatre with all the participants present. During the first five minutes, the participants listened to an oral explanation about the activities to follow and received all the material they needed. Participants in the five stimulated groups (SGs) received the design brief, the external stimulus and instructions, and blank sheets of paper, whilst participants in the baseline group (BG) received the brief without any stimulus. Then, the participants were asked to think of ideas for three minutes without committing any designs to paper (because different participants had different materials and content, this ensured they all had enough time to read all the materials and start developing some ideas). Finally, for the remaining ten minutes, all participants individually generated as many ideas as possible in silence, ideally including both a sketch and a written description for each idea.

4.5 Materials

All participants received the same design problem written on an A4 sheet, as well as blank A4 sheets to sketch and annotate their own ideas. Except for the participants in the baseline group, all participants received one additional sheet with an example solution, i.e. an annotated sketch of a bike (Figure 1).

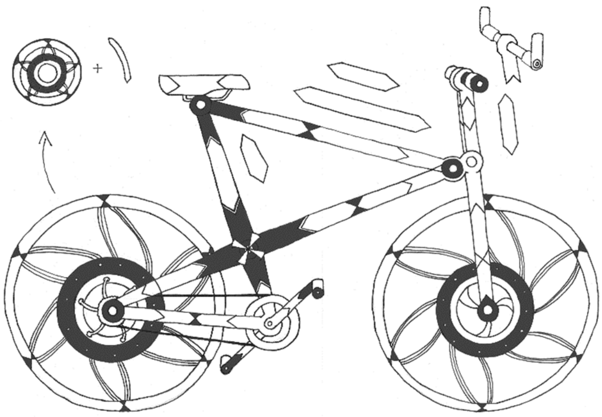

Fig. 1. Example solution provided to participants along with the following description: "*A modular bike with parts of various sizes that can be connected and swapped to fit people with very different heights. Apart from the socketing parts and expansible/contractible wheels, the angles between tubes can also be modified in specific joints*". The sketch used is a modification of the ECO 07 Compactable Urban Bicycle (Aleman, 2009)

The example solution was preceded with the written description: "*Below is an example of how you should present your ideas (i.e. annotated sketches)*". This description was either immediately followed by an instruction regarding the use of features from the examples (discouraging or encouraging) or by no instruction whatsoever. The instructions for the different experimental groups are listed below against a code for each experimental group.

- SG−2 (strictly forbidding): "*make sure you do not use features from this example in your own work*"
- SG−1 (discouraging): "*avoid using features from this example in your own work*"
- SG0 (neutral): no instruction was given
- SG+1 (encouraging): "*consider using features from this example in your own work*"
- SG+2 (strictly requiring): "*make sure you use features from this example in your own work*"

4.6 Analysis

The assessment of the participants' ideas was conducted by three research collaborators, with backgrounds in design research, complexity science, and mechanical engineering. Initially, the three evaluators agreed on the metrics to be included in the analysis. Such metrics should be consistent with most of the existing design fixation literature and should be appropriate for testing our hypothesis.

After that, the three evaluators analysed the design work of a random experimental group together to agree on the assessment method, ultimately reaching a consensus with respect to how to interpret and assess the ideas. Finally, each of the evaluators, who were blind to the experimental groups that they were rating, individually judged a random subset of the remaining ideas. To reduce the chances of mistakes, if any evaluator had trouble judging an idea, this idea was then discussed collectively. We considered "one idea" either to be a sketch or a written description (usually both) that presented an understandable way to solve the problem. Participants often generated more than one idea, but in some cases all ideas could be considered as one, particularly when the idea was a bike. For instance, if the ideas could all be incorporated onto the same bike without interference, then they were considered as a single idea. Conversely, if there were two or more ideas for the same bike component (e.g. frame, wheel, handle bar), then they were considered to be distinct ideas.

The metrics used in the assessment were idea "fluency" and idea "repetition". These metrics are central to many fixation studies (e.g. Jansson & Smith, 1991; Purcell & Gero, 1996; Dahl & Moreau, 2002; Linsey et al., 2010; Cardoso & Badke-Schaub, 2011; Youmans, 2011a, 2011b; Vasconcelos et al., 2017) and they can provide a fairly objective measurement (i.e. a direct count of ideas and features) that is consistent with testing our feature repetition hypothesis. Therefore, other design fixation metrics found in the literature, such as diversity and conceptual distance, were not included into the analysis because they are more subjective (i.e. increased judgement is required on the part of assessors) and are not suited to testing our research hypothesis. Because testing our hypothesis implied identifying and counting ideas and features without any further judgement, we did not require duplicate assessment from our evaluators.

Idea fluency is the total number of ideas generated, also called "quantity" (Shah et al., 2003) or "productivity" (Nijstad et al., 2002) elsewhere. Idea repetition might occur at different levels. For instance, the repetition of idea types, conceptual features, or structural features (although other categorisations and granularity levels might also be identified). With respect to the idea type, we divided the ideas into two broad categories: bike ideas and non-bike ideas, thus by designing a bike the participant would be repeating the idea type. With respect to their conceptual features, we also divided the ideas into two categories: modular ideas or non-modular ideas, thus by generating a modular idea the participant would be repeating the conceptual feature. Finally, we examined the incorporation of structural features in the participants' ideas, in terms of both the number of ideas containing any feature of the example (i.e. is repetition there?) and the number of example features present in each idea (i.e. how much repetition is there?). These features were intentionally included in the example design to permit a measure of fixation inferred because of repetition. There were five structural features: swappable components to change bike size; frame joints (lugs) that act as sockets for the tubes; wheels with bendable spokes; an hourglass-shaped frame; and a saddle that cannot be adjusted in height directly.

Eight participants (4.8%) either did not generate any idea or generated ideas that could not be interpreted by the evaluators; the results from such participants were not included in our analysis. The adjusted number of participants per experimental group is used in the following section.

4.7 Results and discussion

To deal with non-normality in the data, we used Analysis of Variance (ANOVA) with significance values estimated using 1000 bootstrap resamples and planned contrasts (a non-parametric version of the regular ANOVA test). To identify differences between groups, we used five planned contrasts in the analysis. The first contrast compares the baseline group to all stimulated groups, and aims to identify any effect from exposure to the example design (*the fixation test*). The remaining contrasts incorporate only the stimulated groups, and aim to identify any effect from instructions received as follows. The second contrast compares the discouraged and neutral groups (*the discouraging instructions test*); the third compares the encouraged and neutral groups (*the encouraging instructions test*); the fourth compares the two discouraged groups (*the discouraging subtlety test*); and the fifth contrast compares the two encouraged groups (*the encouraging subtlety test*).

Idea fluency

Instructions had no effect on the number of ideas generated. A one-way ANOVA with the total number of ideas (per participant) as the dependent variable showed a significant difference across the groups, $F(5, 157) = 4.37$, $p = .001$, $\eta^2 = .122$. However, planned contrasts revealed that this difference is explained by the baseline group generating on average more ideas than the stimulated groups,

$t(32.7) = 3.64$, $p = .002$, $d = .840$, since no difference was found between the stimulated groups. Table 2 shows summary statistics for these results.

Table 2. Summary of ideas generated per participant across groups

Generated ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of ideas per participant	2.39 (1.31)	1.50 (1.27)	1.30 (0.61)	1.48 (0.49)	1.41 (0.56)	1.54 (1.10)
Range of ideas generated per participant	1-5	1-6	1-3	1-4	1-4	1-6
Total number of ideas	67	39	35	40	38	43
Total number of participants	28	26	27	27	28	28

These results reveal that idea fluency was not influenced by how constraining or encouraging the instructions were. However, the presence of an example design affected idea generation: designing without exposure to stimuli resulted in more ideas being generated, which we interpret as a beneficial isolation from examples (Vasconcelos et al., 2017). This effect is consistent with other studies in which seeing an example caused reduction in the idea fluency (Linsey et al., 2010), although studies have also reported an increase in the idea fluency as a result from external stimulation (Purcell & Gero, 1992) or even no effect at all (Jansson & Smith, 1991). More importantly, our results are different from those found in (Smith et al., 1993), where the number of ideas generated by a baseline group did not exceed that of stimulated groups. However, we should mention that while all groups in our study were asked to present their ideas with both sketches and textual descriptions, the baseline group produced many ideas that were presented only in text. As a result, baseline participants spent less time sketching every idea, which allowed them more time to produce a greater number of less-elaborated ideas, as explained by the "normative representation effect" (Vasconcelos et al., 2016).

Repetition of the idea type

Instructions had no effect on the number of bike ideas generated. A one-way ANOVA with the average number of bike ideas (per participant) as the dependent variable showed a significant difference across the groups, $F(5, 157) = 2.87$, $p = .017$, $\eta^2 = .084$. However, planned contrasts revealed that this difference is explained by participants in the baseline group generating, on average, fewer bike ideas than the stimulated groups, $t(31.4) = -2.59$, $p = .025$, $d = .613$, since no difference was found between the stimulated groups. Table 3 shows summary statistics for these results.

Table 3. Summary of bike and non-bike ideas generated across groups

Bike Ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of bike ideas per participant	1.43 (0.84)	1.27 (0.92)	1.15 (0.36)	1.33 (0.83)	1.33 (0.55)	1.29 (0.53)
Number of bike ideas (and %)	40 (60%)	33 (85%)	31 (89%)	36 (90%)	36 (95%)	36 (84%)
Number of non-bike ideas (and %)	27 (40%)	6 (15%)	4 (11%)	4 (10%)	2 (5%)	7 (16%)

If we adopt a “bike” and “non-bike” categorisation of the ideas generated, these results reveal that the type of idea generated was not influenced by the instructions. This can be seen as a surprising result: by manipulating how participants were instructed to incorporate features of the example, we also anticipated an indirect effect on the type of ideas that they might have generated (because all features were incorporated into a bike idea), but this was not the case. Conversely, whilst the total number of bike ideas was roughly the same for all groups, the proportion of bike ideas generated per participant across all groups was significantly different. In fact, only 60% of all ideas generated by the baseline group were bikes, whereas the stimulated groups had a much greater proportion of bike ideas (89% on average). In summary, bike ideas were likely to be generated irrespective of the experimental condition, but not seeing the example allowed participants from the baseline group to explore different areas of the solution space, again confirming the beneficial isolation effect. Accordingly, design fixation (inferred because of repetition of the idea type) occurred across the stimulated conditions, a result that is broadly consistent with other studies in which seeing an example caused participants to conform to certain types of solutions, thus reducing the breadth of ideas (Jansson & Smith, 1991; Linsey et al., 2010; Cardoso & Badke-Schaub, 2011).

Repetition of the conceptual feature

Instructions had no effect on the number of modular ideas generated. A one-way ANOVA with the average number of modular ideas (per participant) as the dependent variable showed no significant difference across the groups, $F(5, 157) = 1.86$, $p = .104$, $\eta^2 = .056$. Table 4 shows summary statistics for these results.

Table 4. Summary of modular and non-modular ideas generated across groups

Modular ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of modular ideas per participant	0.43 (0.50)	0.08 (0.27)	0.3 (0.47)	0.19 (0.40)	0.19 (0.40)	0.18 (0.39)
Number of modular ideas (and %)	12 (18%)	2 (5%)	8 (23%)	5 (13%)	5 (13%)	5 (12%)
Number of non-modular ideas (and %)	55 (82%)	37 (95%)	27 (77%)	35 (87%)	33 (87%)	38 (88%)

If we adopt a “modular” and “non-modular” categorisation of the ideas, these results reveal that the type of idea generated was not influenced by the instructions. In fact, modular ideas were extremely rare across all groups. Still, the fixation literature tells us that the stimulated groups would generate more modular ideas, an effect similar to the repetition of the idea type (i.e. bikes). In particular, participants who were encouraged to use features from the example in their own work (and modularity was a visible feature on the example provided) did not act accordingly, and participants who were discouraged from copying features generated as many modular ideas as the other groups. The fact that encouraged groups did not produce more modular ideas contradicts previous research (Smith et al., 1993), where results pointed to an increase in feature repetition in encouraged groups. Additionally, previous research has shown that abstract, conceptual features can fixate designers in a similar way to concrete, structural features (Vasconcelos et al., 2017). However, as there was no difference in the proportion of modular ideas between baseline and stimulated groups, we speculate that the idea of modularity (included in the example as a conceptual feature) was not obvious enough to induce fixation effects (inferred because of repetition of the conceptual feature).

Repetition of structural features

Instructions had a significant effect on the number ideas with features of the example provided. A one-way ANOVA with the average number of ideas that contained structural features of the example (per participant) as the dependent variable showed a significant difference across the groups, $F(5, 157) = 9.50$, $p = .000$, $\eta^2 = .232$; a result that was also found for the average number of repeated features per participant, $F(5, 157) = 14.115$, $p = .000$, $\eta^2 = .310$. For the number of ideas containing features of the example solution, planned contrasts revealed a significant difference between the two encouraged and the neutral groups, $t(76.0) = -5.33$, $p = .001$, $d = 1.14$, with the encouraged groups generating more ideas with features from the example. For the average number of features included in the participants’ ideas, planned contrasts showed a marginally significant difference between the baseline group and all stimulated groups, $t(44.6) = -1.79$, $p = .086$, $d = .338$, and again, between the two encouraged and the neutral groups, $t(71.0) = -6.47$, $p = .001$, $d = 1.29$, with the stimulated groups repeating on average more features than the baseline condition due to an increased feature repetition of the encouraged groups. Table 5 shows summary statistics for these results.

Table 5. Summary of modular features incorporated into the participants' ideas and ideas with modular features across groups

Feature repetition	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of features incorporated per idea	0.36 (0.62)	0.12 (0.33)	0.30 (0.47)	0.15 (0.36)	1.00 (1.14)	1.43 (1.07)
Range of features incorporated per participant	0-2	0-1	0-1	0-1	0-3	0-3
Total number of ideas with features incorporated (and %)	8 (12%)	3 (8%)	8 (21%)	4 (11%)	16 (44%)	24 (57%)

These results reveal that encouraging instructions influenced the incorporation of structural features from the example. On average, participants in encouraged groups incorporated more example features per idea and generated more ideas that incorporated such features. Additionally, participants in discouraged groups produced results similar to the neutral and baseline groups. This result is inconsistent with previous research from Chrysikou and Weisberg (2005), as constraining instructions did not decrease repetition, but consistent with Smith et al. (1993), as encouraging instructions did increase repetition. However, contrary to the results from Smith et al. (1993), our discouraged (but still stimulated) groups did not significantly repeat more features than a baseline group, contradicting the idea that participants could not forget the example they had seen. It seems that a few structural features were naturally likely to be incorporated into the ideas generated (supported by the similar results from the non-encouraged groups), with encouraged groups incorporating more features because they were instructed to do so. When comparing the results from the repetition of structural features to the repetition of the conceptual feature used in this study, we can infer that concrete structural features are more easily copied than abstract conceptual features from examples (similar to what has been suggested in previous studies (Zahner et al., 2010; Cheong et al., 2014; Feng et al., 2014)). This is despite the opportunity that conceptual features offer for inspiring solutions across multiple categories of solution (e.g. products, services, etc.). Alternatively, it is possible that either the conceptual feature chosen or its incorporation into the example were not effective and thus not fully comprehended by the participants. Again, we should highlight that the baseline group produced more ideas but proportionally less sketches than the stimulated groups. This puts the stimulated groups in an unfavourable position with respect to the count of structural features incorporated, since these groups had to represent a shape of the bike in which repetition could be more easily recognised – it is difficult to identify structural repetition when the idea is represented only by text – thus possibly biasing the results.

4.8 General discussion and limitations

In summary, except for the repetition of structural features, the instructions manipulation had no effect on participants' idea generation according to the creativity metrics used in this study. Idea fluency and the repetition of the idea type were only affected by the exposure to the example design, and the repetition of the conceptual feature did not seem to be affected by either instructions or exposure to the example. This last set of results conflicts with a previous study that found that a pictorial representation of modularity (as a conceptual and abstract feature) induced fixation effects in a similar way to a more concrete bike example (Vasconcelos et al., 2017). Additionally, looking more carefully into our data reveals that the generation of modular ideas happened quite rarely and randomly across groups, and it is possible that the feature "modularity", in particular, failed to inspire participants. These findings call for further investigation and extension of this study, potentially by incorporating other stimuli. As such, to test whether our previous findings are robust to a different stimulus with different structural and conceptual features (while keeping all other experimental variables the same), we now report an additional analysis done on a second dataset. Additionally, in Study 1 we did not check the inter-rater agreement, a test which is often performed for similar studies to demonstrate the reliability of the evaluators' subjective assessment when working individually. The assessment performed was relatively objective and yet involved many interactions between evaluators (thus we expected a high agreement between them and we believe that such interaction allowed a reliable analysis of the idea). However, because there was no redundancy in the assessment, we cannot quantify how good this agreement might have been. As such, in Study 2 we now compute and report agreement coefficients.

5. STUDY 2

5.1 Experimental method, participants, task and problem, and procedure

Studies 1 and 2 share the same experimental method and drew participants from the same population. One hundred and thirty-five participants were assigned to five experimental groups (*n* = 26). The task, design problem, and procedure were the same as in Study 1, except for one aspect: we did not have a new baseline condition in Study 2 since we already had data from a non-stimulated population to which we could compare our new stimulated groups.

5.2 Materials

To allow a neat comparison of Studies 1 and 2, we chose a similar example solution to be given to the participants. Whilst the problem could be solved in many ways and with several underlying properties being applied, it was expected that developing an "adaptable" solution would be a common approach in this context. Such an approach could equally be realized by invoking a modular concept (e.g. by interchangeable parts) or a telescopic concept (e.g. by extendable parts). As a result, we opted for using a telescopic bike (Figure 2) with both sketch and textual description of a similar quality to that used in Study 1 to avoid influencing the participants' perception of the ideas due to different sketch quality (Kudrowitz et al., 2012). The description of the example solution was the same used in Study 1 ("*Below is an example of how you should present your ideas (i.e. annotated sketches)*"), and it was followed by the same set of discouraging, encouraging, or neutral instructions regarding the use of features from the example.

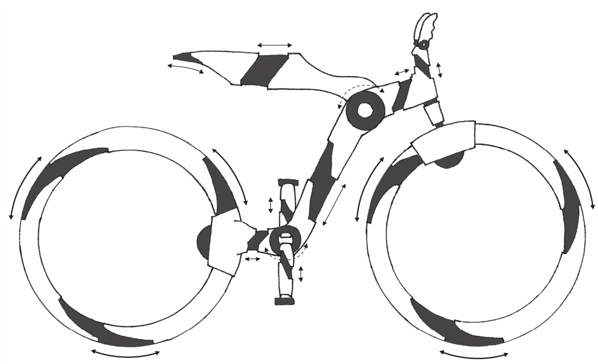

Fig. 2. Example solution provided to participants along with the following description: "*A telescoping bike with parts that can be extended or shortened to fit people with very different heights. Apart from the adjustable tubes and wheels, the angles between tubes can also be modified in specific joints*". The sketch used is a modification of the Zee-K Ergonomic Bike (Floss, 2010)

5.3 Analysis

The assessment of the participants' ideas was conducted by the first two authors of this study, with backgrounds in design research and experimental psychology respectively. First, the evaluators discussed the metrics of the analysis and assessed a random sample of ideas to reach a satisfactory agreement. As in Study 1, the coding scheme used here was not very subjective – the evaluators spotted the corresponding features in the drawings and reported them. However, the drawings were not always clear and the handwriting was often poor. Consequently, the assessment was prone to human error when interpreting what was contained in each idea, and differences between evaluators were more likely to be attributed to that, instead of divergence in individual opinions. As such, the two evaluators assessed the complete set of ideas individually to reveal errors. Such errors were identified by reviewing all ratings and looking for ideas with discrepancies in many variables, which often indicated ideas with a poor representation. After clarifying such cases, we computed Cohen's Kappa coefficients for each metric, as shown in Table 6. When considering the high agreement coefficients obtained here and the very brief training provided to the evaluators, it is clear that there was little subjectivity when assessing the ideas and similar coefficients should be expected for Study 1.

Whilst most particularities of the assessment where the same as in Study 1, including the metrics used (i.e. idea fluency and idea repetition), the change in the stimulus required a new set of structural features to be present in the example design. Again, there were five features intentionally incorporated into the example: extendable components to change bike size; frame joints (lugs) that rotate to adjust angles between parts; wheels with no spokes; an open-shaped frame; and a cantilevered saddle that can be adjusted directly.

Table 6. Ideation metrics used in this study and their computed agreement coefficients and interpretation

Metric	**Cohen's Kappa coefficient**	**Interpretation (Landis & Koch, 1977)**
A bike idea	.927	Almost perfect
A telescopic idea	.893	Almost perfect
A modular idea	.877	Almost perfect
Extendable components	.823	Almost perfect
Rotating joints	.768	Substantial
No spokes	.769	Substantial
Open frame	.915	Almost perfect
Cantilevered saddle	.852	Almost perfect

Five participants (3.7%) either did not generate any idea or generated ideas that could not be interpreted by the evaluators; the results from such participants were not included in our analysis. The adjusted number of participants per experimental group is used in the following section.

5.4 Results and discussion

Idea fluency

Instructions had no effect on the number of ideas generated. A one-way ANOVA with the total number of ideas (per participant) as the dependent variable showed a significant difference across the groups, $F(5, 152) = 4.79$, $p = .000$, $\eta^2 = .136$. However, planned contrasts revealed that this difference is explained by the baseline group generating on average more ideas than the stimulated groups, $t(32.4) = 3.59$, $p = .002$, $d = .844$, since no difference was found between the stimulated groups (Table 7 shows summary statistics for these results). This set of results matches very closely those obtained in Study 1, confirming that the manipulation of instructions had no effect on the number of ideas generated, but exposure to an example design did.

Table 7. Summary of ideas generated per participant across groups

Generated ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of ideas per participant	2.39 (1.31)	1.29 (0.46)	1.56 (1.45)	1.44 (0.65)	1.38 (0.64)	1.62 (0.85)
Range of ideas generated per participant	1-5	1-2	1-8	1-3	1-3	1-4
Total number of ideas	67	36	39	36	36	42
Total number of participants	28	28	25	25	26	26

Repetition of the idea type

Instructions had no effect on the number of bike ideas generated. A one-way ANOVA with the average number of bike ideas (per participant) as the dependent variable showed a significant difference across the groups, $F(5, 152) = 2.44$, $p = .037$, $\eta^2 = .074$. However, planned contrasts revealed that this difference is explained by the baseline group generating, on average, fewer bike ideas than the stimulated groups, $t(31.4) = -2.45$, $p = .029$, $d = .576$, since no difference was found between the stimulated groups (Table 8 shows summary statistics for these results). As in Study 1, the manipulation of instructions had no effect on the number of bike ideas generated, even though the instructions had explicit directions towards the use of features from the bike example and whether participants should comply with the instructions, we expected to observe an indirect effect on the number of bike ideas being generated.

Table 8. Summary of bike and non-bike ideas generated across groups

Bike Ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of bike ideas per participant	1.43 (0.84)	0.929 (0.22)	0.962 (0.14)	0.867 (0.33)	0.904 (0.25)	0.936 (0.18)
Number of bike ideas (and %)	40 (60%)	33 (92%)	33 (85%)	31 (86%)	30 (84%)	37 (88%)
Number of non-bike ideas (and %)	27 (40%)	3 (8%)	6 (15%)	5 (14%)	6 (16%)	5 (12%)

Repetition of the conceptual feature

Instructions had an effect on the number of telescopic ideas generated. A one-way ANOVA with the average number of telescopic ideas (per participant) as the dependent variable showed no significant difference across the groups, $F(5, 152) = 1.75$, $p = .126$, $\eta^2 = .055$. However, planned contrasts revealed a significant difference between the two encouraged groups and the neutral group, $t(152) = -2.27$, $p = .021$, $d = .560$, with the encouraged groups generating more telescopic ideas (Table 9 shows summary statistics for these results).

Table 9. Summary of telescopic and non-telescopic ideas generated across groups

Telescopic ideas	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number of telescopic ideas per participant	0.43 (0.50)	0.08 (0.27)	0.3 (0.47)	0.19 (0.40)	0.19 (0.40)	0.18 (0.39)
Number of telescopic ideas (and %)	19 (28%)	16 (44%)	7 (18%)	8 (22%)	15 (42%)	19 (45%)
Number of non-telescopic ideas (and %)	48 (72%)	20 (56%)	32 (82%)	28 (78%)	21 (58%)	23 (55%)

If we adopt a “telescopic” and “non-telescopic” categorisation of the ideas, these results reveal that the type of idea generated was influenced by the instructions, a finding that differs from that reported in Study 1. As discussed before, it is possible that the notion of extendibility (included in the example as a conceptual feature) was more recognisable than modularity, thus being more easily copied. Interestingly, if we look at the number of *modular* ideas in Study 2, a similar contrast between the two encouraged and the neutral groups showed a marginally significant result, $t(41.0) = 2.02$, $p = .053$, $d = .507$, but now with the encouraged groups generating less modular ideas (encouraged, M = .162, SD = .334; neutral, M = .347, SD = .394). A possible explanation for this is that, while those participants who were encouraged to reproduce features from the telescopic example increased the generation of telescopic ideas, they were also diverted or pushed away from the generation of modular ideas (these two approaches were so commonly adopted that together they represented 46% of the ideas from the baseline group). Whereas research has already reported the occurrence of such phenomena (Vasconcelos et al., 2017), this last set of results should be interpreted with caution since they are derived after the formal analysis had been made, rather than being hypothesised and planned prior to the analysis.

Repetition of structural features

Instructions had a marginally significant effect on the number ideas with features of the example provided. A one-way ANOVA with the average number of ideas that contained structural features of the example (per participant) as the dependent variable showed a marginally significant difference across the groups, $F(5, 152) = 2.01$, $p = .081$, $\eta^2 = .062$, and a significant difference for the average number of repeated features per participant, $F(5, 152) = 2.73$, $p = .021$, $\eta^2 = .082$. Planned contrasts revealed significant differences between the encouraged groups and the neutral group for both the number of ideas containing features of the example solution, $t(152) = -2.67$, $p = .004$, d = .710, and the number of features included in the participants’ ideas, $t(152) = -3.04$, $p = .004$, $d = .736$, with the encouraged groups generating more ideas with these features and repeating more of these features per idea (Table 10 shows summary statistics for these results).

Table 10. Summary of telescopic features incorporated into the participants' ideas and ideas with telescopic features across groups

Feature repetition	**BG**	**SG−2**	**SG−1**	**SG0**	**SG+1**	**SG+2**
Mean (and SD) for the number features incorporated per idea	1.21 (1.03)	1.43 (1.55)	1.00 (1.19)	0.84 (1.07)	1.65 (1.60)	2.08 (1.72)
Range of features incorporated per participant	0-4	0-6	0-4	0-4	0-8	0-5
Total number of ideas with features incorporated (and %)	24 (36%)	23 (64%)	15 (38%)	13 (36%)	23 (64%)	26 (62%)

As in Study 1, these results show that only the encouraging instructions influenced the incorporation of structural features from the example. On average, participants in encouraged groups incorporated more example features per idea and generated more ideas that incorporated such features, whereas discouraged groups produced results similar to the neutral and baseline groups. Whilst these findings are consistent with those of Study 1, here we observed considerably smaller F-statistics and effect sizes. This can be attributed to the baseline group generating more ideas with features of the telescopic example than of the modular one, $t(54) = -3.59$, $p = .001$, $d = .960$, and incorporating more of such features into the ideas when compared to the incorporation of modular features, $t(44.3) = -3.77$, $p = .000$, $d = 1.01$. Indeed, there was a general tendency of increased feature repetition in the stimulated groups of Study 2, so we can assume that the features of the telescopic example were more likely to appear in the participants' ideas in Study 2 than the modular features in Study 1.

5.5 General discussion and limitations

Study 2 results replicate most of what we have found in Study 1. In fact, the statistics regarding fluency and the repetition of the idea type were almost identical, indicating that these findings are robust to different stimuli. With respect to the conceptual feature, even though we have found some repetition in the encouraged groups, participants in Study 2 generated on average less telescopic ideas (M = .407, SD = .460, $t(263) = 3.02$, $p = .003$, $d = .370$) and more modular ideas (M = .262, SD = .396, $t(247.4) = -2.79$, $p = .006$, $d = .344$) than Study 1 participants (telescopic ideas, M = .575, SD = .448; modular ideas, M = .138, SD = .318). This is a surprising result, as it shows an apparent resistance towards incorporating the underlying mechanisms behind the adaptable bike to which participants were exposed. Moreover, it might be the case that this resistance was a conscious choice, as though participants deliberately tried not to conform to whatever they were shown, either because the example already pre-exhausted the solution space they could have explored (Perttula & Liikkanen, 2006), or because the participants were suspicious or exhibited demand awareness (Page, 1981). Figure 3 illustrates a selection of participants' ideas that were bikes, not bikes, modular, and telescopic.

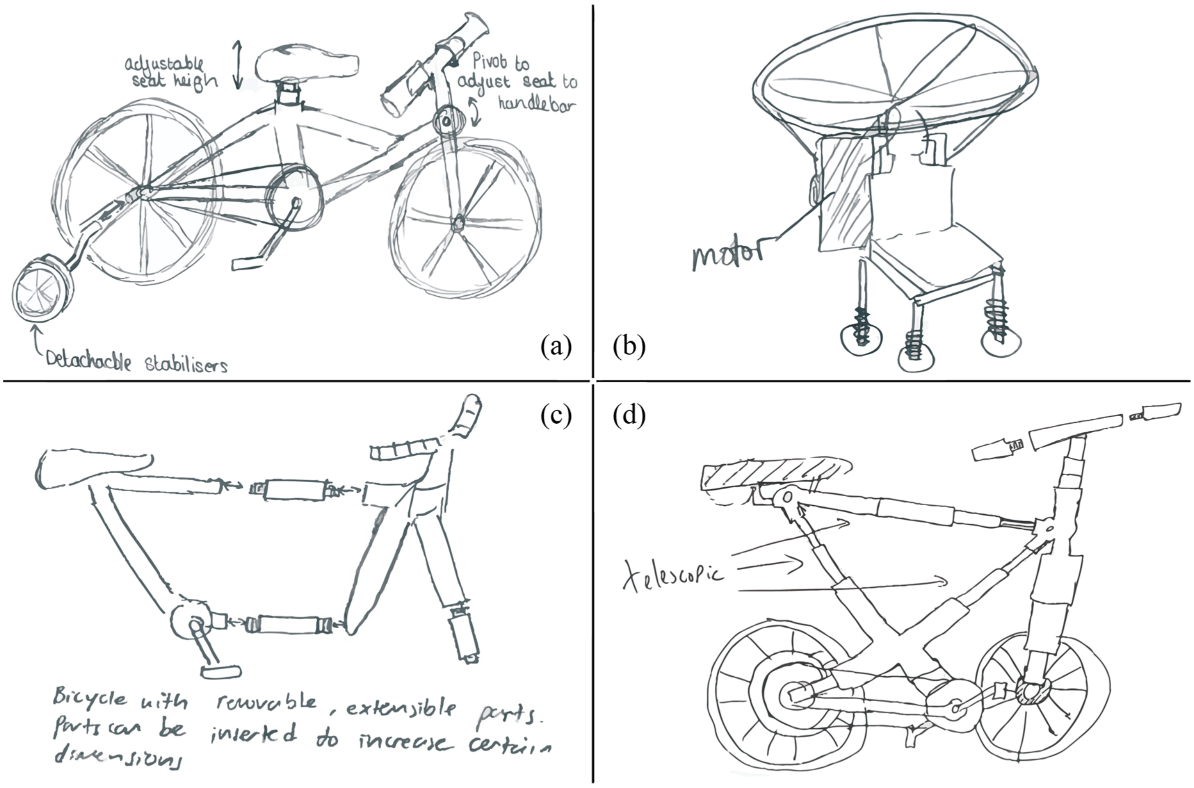

Fig. 3. Some of the participants' ideas: (a) bike idea, (b) non-bike idea, (c) modular idea, and (d) telescopic idea. Note that these examples are not representative of the design work of any particular experimental group

The results for structural feature repetition provide further support for the effectiveness of encouraging instructions in inspiration and fixation experiments, as shown in Study 1 and reported in previous research (Smith et al., 1993). However, the data also revealed a critical aspect of fixation studies: idea repetition results will depend on the kind of feature that experimenters incorporate into the analysis. Presumably, differences in feature repetition results between stimulated and non-stimulated participants are more pronounced when those features represent rare, unusual solutions to the design problem, thus being very unlikely to emerge during idea generation in non-stimulated groups. Conversely, repetition results might be similar if the features analysed represent obvious solutions, thus being highly likely to appear in ideas with or without external stimulation (although there is recent evidence to the contrary (Viswanathan et al., 2016)). Whereas the features used in both our studies are easily comparable (i.e. following a similar structure and referring to the same bike parts), we believe that the features used in Study 2 were more common than those in Study 1, especially the "*extendable components to change bike size*", which frequently appeared in the designs of all groups. This might have biased the idea repetition analysis and therefore compromised the comparability of the two studies in this respect (see Figure 4 for a comparison of feature repetition results between the two studies).

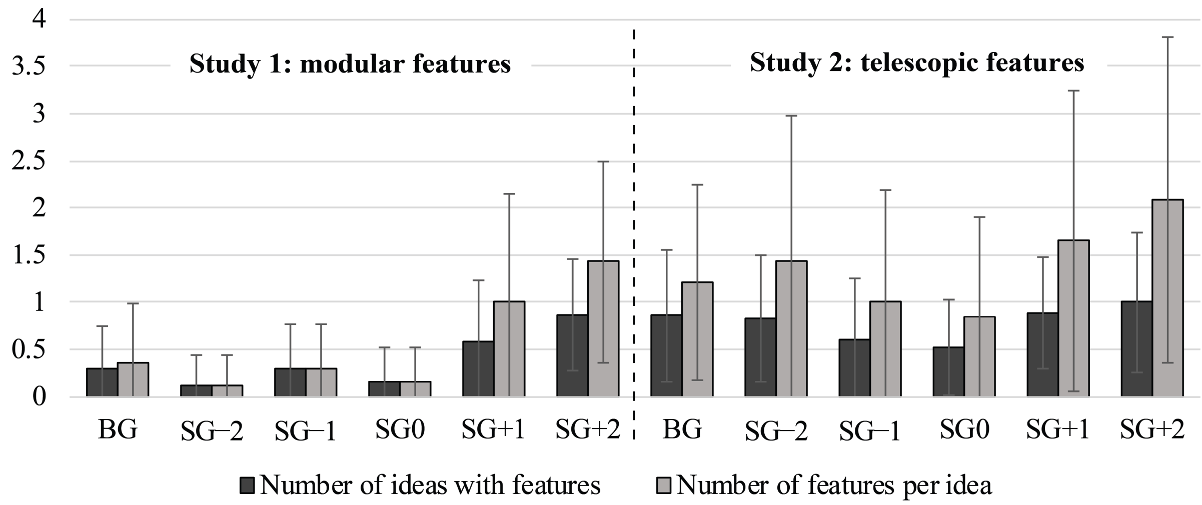

Fig. 4. Bar chart showing the average number of ideas with features of the example and average number of example features per idea across groups (averages given per participant; error bars indicate standard deviation)

The main limitations of the studies reported here are the duration of the generation session, the pool of participants chosen, and the design problem used. The idea generation session in this study was ten minutes long, which can be considered short when compared to other fixation studies in which generation sessions typically lasted for thirty or sixty minutes (Vasconcelos & Crilly, 2016). Also, research suggests that novel ideas tend to occur later in the idea generation session (Kudrowitz & Dippo, 2013), thus the short session adopted for this study might have contributed to inflated fixation scores (when considering fluency and the repetition of the respective idea types). However, it is also possible that ten minutes of ideation might have been too short for fixation to take place (or to be measured at least), although we believe that this is unlikely, as research has already shown that having extra time helps to diminish fixation effects (Tseng et al., 2008). The participants in our studies were undergraduate students and the generation session was part of an ongoing engineering course. This might have resulted in a more diligent participant behaviour when compared to other studies in which participants and experimenters did not have a student-lecturer relationship. As a result, the setup adopted for this study might have contributed to the extent to which participants adhered to the instructions. Additionally, the behaviour of design students (or novices) and practitioners (or experts) differs, both with respect to the design process in general (for a review see Dinar et al., 2015), and design fixation in particular (for a review see Vasconcelos & Crilly, 2016). Thus, our results might have been different had our participants been more experienced. The design problem used in this study was chosen partly because it was unlikely that the participants had designed solutions to it before. As such, it is possible that using familiar problems will produce different results, although research has demonstrated that fixation effects can be observed with both familiar and unfamiliar problems (Jansson & Smith, 1991).

Finally, although it is not necessarily a limitation of this type of work, but a characteristic of it, the evaluation method used in these studies can be considered subjective to some extent (yet, consensual assessment is still the standard method in idea generation and design fixation studies (see Vasconcelos & Crilly, 2016)). Whilst new methods are available today, such as eye-tracking and neuro-imaging they either require additional techniques for triangulation or are still very recent and their potential for providing more objective, meaningful data is yet to be developed (Editorial board of IJDCI, 2013). With the aim to achieve higher objectivity levels when analysing data from idea generation experiments, two recent studies had their participants designing in a digital environment, such as computer games (Neroni et al., 2017) and computer-aided design tools (Zhang et al., 2017). Apart from allowing more objective and comparable analyses of the design activity, these approaches may enable a series of benefits, such as: using design problems and materials that are not easily manipulated in classroom settings, unobtrusive data capture throughout the design task, compatibility with other digital tools for data capture, and better control over the format and quality of the output generated. These are exciting opportunities not just for future work in design fixation; they show themselves as new methodological paradigms to study human cognition more broadly.

6. CONCLUSION AND FUTURE WORK

We have tested the influence of instructions on idea generation. More specifically, we analysed how discouraging, encouraging, and neutral instructions may affect the number of ideas generated and the incorporation of the example or its parts into the participants' ideas. The instructions used were provided along with an external stimulus and its description. It is important to differentiate the descriptions from instructions because in this study we have controlled the former but manipulated the latter. We found that instructions shaped the idea generation of our participants to some extent. When encouraged or required to use features from the example, participants consistently copied structural features but some failed to copy a more abstract conceptual feature. When discouraged or forbidden from using features from the example however, most participants failed to reduce the number of features copied. This result allows us to infer that more concrete features are easier to recognise – and thus reproduce – than more abstract features such as modularity. However, this may reflect a peculiarity of the modularity feature, since an equivalent approach, such as extendibility, produced slightly different results. Additionally, the results might indicate that positive instructions are more effective than negative ones, irrespective of how flexible or strict the instructions are. This can tell us how to frame future instructions, whether that is with respect to experimental stimuli in research, illustrative cases in design education, and inspirational examples in design practice.

Regardless of the instructions given to the participants, using different conceptual and structural features between studies allowed us to explore how two alternative but equivalent approaches to solve a design problem can compete for attention. When participants complied with the encouraging instructions and generated more telescopic ideas, they also decreased the number of modular ideas they generated – we call this *design diversion*. Two other observations were made when analysing the datasets. The first is that the telescopic bike conditions generated more modular ideas and less telescopic ideas than the modular bike conditions, and we interpret this as the participants deliberately exploring an alternative solution space. The second is that feature repetition calculations can be misleading depending on how common or obvious the features might be (e.g. the bike has a saddle *vs.* the bike has wings), but also on their level of granularity (e.g. the idea is a transportation mode, it is a bike, it is a modular bike, the pedals are made of modular-blocks, and so on), with repetition results likely varying according to each level. As such, future studies should reflect on both the commonness and the granularity level of the features under consideration, and could report results according to these categories of features (e.g. a quadrant chart of features, having commonness and granularity as two orthogonal continuums). Nevertheless, all these potential findings need support from more data and are thus open for future investigation.

Again, irrespective of how discouraging the instructions were, stimulated participants still exhibited fixation effects due to their exposure to the example design (i.e. they generated fewer ideas and proportionally more bike ideas than the baseline group). This result is in line with many other design fixation experiments in which participants become stuck on a specific idea type. However, it is important to emphasise that in this study the description of the stimulus itself could also be causing fixation as the stimulus was presented to the participants as "*an example of how they should present their ideas*". Thus, perhaps there is an implicit suggestion for the participants to produce ideas that are similar to the example, i.e. a bike. Another possible explanation is that in inspiration and fixation studies the examples overwhelm the participants' interpretation of the problem, to the extent that the examples themselves become part of the participants' problem representation. Future studies could investigate such possibilities and complement our understanding about stimuli introduction by manipulating the descriptions provided to designers and by testing how examples interfere with problem representation. It is also important to consider here and in similar future studies, that copying and repeating features from examples does not necessarily imply fixation. Fixation might be inferred when participants are discouraged from copying (or maybe when they are given no guidance at all), and yet they still copy, otherwise they may only be complying with the descriptions and instructions in the experiments. Similarly, we should be cautious when referring to "idea repetition" for those participants who are given no stimulus but spontaneously generate ideas that incorporate features that also happen to be in the example provided to others.

Finally, when considering design practice, design teams and project managers can benefit from being aware of the design fixation knowledge and from the findings discussed in this paper, applying such knowledge depending on the scope and stage of the design project. For instance, the idea of design diversion might be considered by project managers for strategically removing complementary areas of the solution space that are not of interest. Also, the results about the efficacy of instructions and the abstraction of stimuli reported here are relevant to how inspirational stimuli should be framed when presented to designers. This is particularly important for the development and implementation of computer-aided design tools that provide designers with external stimuli. Much has been researched on how such software tools might be structured and interacted with, and what form the inspirational stimuli should take (Shneiderman, 2000; Töre Yargin & Crilly, 2015). However, it is also important to understand how those stimuli should be introduced, whether by description, instruction, or both. Our findings show that using positive instructions (either requiring or encouraging) in inspiration tools can guarantee that designers will incorporate features from a given stimulus, and that choosing a concrete stimulus will increase feature incorporation even further. However, should designers be steered towards or away from the repetition of structural features, directed to identify conceptual features, or simply be left alone to interpret and respond as they see fit? By developing a better understanding of how stimuli instructions influence idea generation, we will move closer to answering such questions and thereby become more capable of supporting creative design.

ACKNOWLEDGEMENTS

First of all, out with Temer! This work was supported by the CAPES Foundation, Ministry of Education of Brazil [grant number BEX 11468/13-0] and the UK's Engineering and Physical Sciences Research Council [grant number EP/K008196/1]. Research data supporting this publication are available from the University of Cambridge data repository: https://doi.org/10.17863/CAM.7571

REFERENCES

Aleman, V. (2009). ECO 07 - Compactable Urban Bicycle, accessed February 10, 2017, available at https://www.behance.net/gallery/293563/ECO-07-Compactable-Urban-Bicycle.

Atilola, O., & Linsey, J. (2015). Representing analogies to influence fixation and creativity: A study comparing computer-aided design, photographs, and sketches. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *29*(02), 161–171.

Cardoso, C., & Badke-Schaub, P. (2011). The influence of different pictorial representations during idea generation. *The Journal of Creative Behavior*, *45*(2), 130–146.

Chakrabarti, A., Sarkar, P., Leelavathamma, B., & Nataraju, B. S. (2005). A functional representation for aiding biomimetic and artificial inspiration of new ideas. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *19*(2), 113–132.

Cheng, P., Mugge, R., & Schoormans, J. P. L. (2014). A new strategy to reduce design fixation: Presenting partial photographs to designers. *Design Studies*, *35*(4), 374–391.

Cheong, H., Hallihan, G., & Shu, L. H. (2014). Understanding Analogical Reasoning in Biomimetic Design: An Inductive Approach. In J. S. Gero (Ed.), *Design Computing and Cognition '12*. Dordrecht: Springer Netherlands, 21–39.

Chrysikou, E. G., & Weisberg, R. W. (2005). Following the Wrong Footsteps: Fixation Effects of Pictorial Examples in a Design Problem-Solving Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 1134–1148.

Dahl, D. W., & Moreau, P. (2002). The Influence and Value of Analogical Thinking during New Product Ideation. *Journal of Marketing Research*, *39*(1), 47–60.

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical studies of designer thinking: past, present, and future. *Journal of Mechanical Design, 137*(2), 021101.

Eckert, C. M. (1997). Design inspiration and design performance. *Proceedings of the 78th World Conference of the Textile Institute*, Thessaloniki, Greece, 369–387.

Editorial board of IJDCI. (2013). Perspectives on design creativity and innovation research. *International Journal of Design Creativity and Innovation, 1*(1), 1–42.

Feng, T., Cheong, H., & Shu, L. H. (2014). Effects of Abstraction on Selecting Relevant Biological Phenomena for Biomimetic Design. *Journal of Mechanical Design*, *136*(11), 111111.

Floss, G. H. (2010). Zee-K Ergonomic Bike, accessed February 10, 2017, available at http://www.coroflot.com/gabrielfloss/zee-k-ergonomic-bike1.

Fu, K., Cagan, J., & Kotovsky, K. (2010). Design Team Convergence: The Influence of Example Solution Quality. *Journal of Mechanical Design*, *132*(11), 111005.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2012). Find your inspiration: exploring different levels of abstraction in textual stimuli. *2nd international conference on design creativity (ICDC2012), Glasgow UK*.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2014). What inspires designers? Preferences on inspirational approaches during idea generation. *Design Studies*, *35*(1), 29–53.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11.

Kudrowitz, B., & Dippo, C. (2013). Getting to the novel ideas: exploring the alternative uses test of divergent thinking. *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers, V005T06A013–V005T06A013.

Kudrowitz, B., Te, P., & Wallace, D. (2012). The influence of sketch quality on perception of product-idea creativity. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *26*(03), 267–279.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159.

LeFevre, J.-A., & Dixon, P. (1986). Do Written Instructions Need Examples? *Cognition and Instruction*, *3*(1), 1–30.

Liikkanen, L. A., & Perttula, M. (2008). Inspiring design idea generation: insights from a memory-search perspective. *Journal of Engineering Design*, *21*(5), 545–560.

Linsey, J., Wood, K. L., & Markman, A. B. (2008). Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *22*(02), 85–100.

Linsey, J., Tseng, I., Fu, K., Cagan, J., Wood, K., & Schunn, C. (2010). A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. *Journal of Mechanical Design*, *132*(4), 041003.

Linsey, J., & Wood, K. (2007). Wordtrees: a method for design-by-analogy. In: *ASEE annual Conference*, AC 2008-1669.

Lujun, Z. (2011). Design fixation and solution quality under exposure to example solution. *Computing, Control and Industrial Engineering (CCIE), 2011 IEEE 2nd International Conference on* (Vol. 1). IEEE, 129–132.

Neroni, M. A., Vasconcelos, L. A., & Crilly, N. (2017). Computer-Based "Mental Set" Tasks: An Alternative Approach to Studying Design Fixation. *Journal of Mechanical Design, 139*(7), 071102.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, *38*(6), 535–544.

Page, M. M. (1981). Demand Compliance in Laboratory Experiments. In J. T. Tedeschi (Ed.), *Impression Management Theory and Social Psychological Research*. Academic Press, 57–82.

Perttula, M., & Liikkanen, L. A. (2006). Exposure effects in design idea generation: unconscious conformity or a product of sampling probability. *Proceedings of NordDesign*, Reykjavik, Iceland: The Design Society, 42–55.

Perttula, M., & Sipilä, P. (2007). The idea exposure paradigm in design idea generation. *Journal of Engineering Design*, *18*(1), 93–102.

Purcell, A. T., & Gero, J. S. (1992). Effects of Examples on the Results of a Design Activity. *Knowledge-Based Systems*, *5*(1), 82–91.

Purcell, A. T., & Gero, J. S. (1996). Design and other types of fixation. *Design Studies*, *17*(4), 363–383.

Sarkar, P., & Chakrabarti, A. (2008). The effect of representation of triggers on design outcomes. *AI Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *22*(02), 101–116.

Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Design Studies*, *24*(2), 111–134.

Shneiderman, B. (2000). Creating Creativity: User Interfaces for Supporting Innovation. *ACM Transactions on Computer-Human Interaction*, *7*(1), 114–138.

Siangliulue, P., Chan, J., Gajos, K. Z., & Dow, S. P. (2015). Providing Timely Examples Improves the Quantity and Quality of Generated Ideas. *ACM SIGCHI Conference on Creativity and Cognition*., Glasgow, United Kingdom: ACM Press, 83–92.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, *21*(6), 837–845.

Töre Yargin, G., & Crilly, N. (2015). Information and interaction requirements for software tools supporting analogical design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *29*(02), 203–214.

Tsenn, J., Atilola, O., McAdams, D. A., & Linsey, J. S. (2014). The effects of time and incubation on design concept generation. *Design Studies*, *35*(5), 500–526.

Vasconcelos, L. A., Neroni, M. A., & Crilly, N. (2016). Fluency results in design fixation experiments: an additional explanation. *The 4th International Conference on Design Creativity*. Atlanta, GA: The Design Society.

Vasconcelos, L. A., Cardoso, C. C., Sääksjärvi, M., Chen, C.-C., & Crilly, N. (2017). Inspiration and Fixation: The Influences of Example Designs and System Properties in Idea Generation. *Journal of Mechanical Design*, *139*(3), 031101–031101-13.

Vasconcelos, L. A., & Crilly, N. (2016). Inspiration and fixation: Questions, methods, findings, and challenges. *Design Studies*, *42*, 1–32.

Vattam, S., Wiltgen, B., Helms, M., Goel, A. K., & Yen, J. (2011). DANE: Fostering Creativity in and through Biologically Inspired Design. In T. Taura & Y. Nagai (Eds.), *Design Creativity 2010*. Springer London, 115–122.

Viswanathan, V., Atilola, O., Esposito, N., & Linsey, J. (2014). A study on the role of physical models in the mitigation of design fixation. *Journal of Engineering Design*, *25*(1–3), 25–43.

Viswanathan, V., Tomko, M., & Linsey, J. (2016). A study on the effects of example familiarity and modality on design fixation. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *30*(02), 171–184.

Yilmaz, S., Seifert, C. M., & Gonzalez, R. (2010). Cognitive heuristics in design: Instructional strategies to increase creativity in idea generation. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *24*(3), 335–355.

Youmans, R. J. (2011a). Design fixation in the wild: Design environments and their influence on fixation. *The Journal of Creative Behavior*, *45*(2), 101–107.

Youmans, R. J. (2011b). The effects of physical prototyping and group work on the reduction of design fixation. *Design Studies*, *32*(2), 115–138.

Youmans, R. J., & Arciszewski, T. (2014). Design fixation: Classifications and modern methods of prevention. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *28*(2), 129–137.

Zahner, D., Nickerson, J. V., Tversky, B., Corter, J. E., & Ma, J. (2010). A fix for fixation? Rerepresenting and abstracting as creative processes in the design of information systems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *24*(02), 231.

Zhang, H. Z., Xie, C., & Nourian, S. (2017). Are their designs iterative or fixated? Investigating design patterns from student digital footprints in computer-aided design software. *International Journal of Technology and Design Education*, 1–23.

Zhao, M. (2013). Seek it or let it come: how designers achieve inspirations. *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2779–2784.

FIGURES

Fig. 1. Example solution provided to participants along with the following description: "*A modular bike with parts of various sizes that can be connected and swapped to fit people with very different heights. Apart from the socketing parts and expansible/contractible wheels, the angles between tubes can also be modified in specific joints*". The sketch used is a modification of the ECO 07 Compactable Urban Bicycle (Aleman, 2009)

Fig. 2. Example solution provided to participants along with the following description: "*A telescoping bike with parts that can be extended or shortened to fit people with very different heights. Apart from the adjustable tubes and wheels, the angles between tubes can also be modified in specific joints*". The sketch used is a modification of the Zee-K Ergonomic Bike (Floss, 2010)

Fig. 3. Some of the participants' ideas: (a) bike idea, (b) non-bike idea, (c) modular idea, and (d) telescopic idea. Note that these examples are not representative of the design work of any particular experimental group

Fig. 4. Bar chart showing the average number of ideas with features of the example and average number of example features per idea across groups (averages given per participant; error bars indicate standard deviation)

AUTHOR BIOGRAPHIES

Luis A. Vasconcelos is a PhD Student in the Design Practice Group at the Cambridge Engineering Design Centre. With a background in design research and methodology, his research explores new paths in design inspiration and fixation studies. He is interested in the design process at both macro and micro levels.

Maria A. Neroni is a Research Associate at the Design Practice Group in the Cambridge Engineering Design Centre. With a background in Experimental Psychology, Cognitive Neuroscience and Education, her current research focuses on developing novel experimental paradigms to investigate creativity in engineering design.

Dr Nathan Crilly is a Senior Lecturer in Engineering Design at the University of Cambridge. He leads the Design Practice research group at the Engineering Design Centre. His research interests lie in design, creativity and communication. He and his research group employ an interdisciplinary approach to studying the development of products, the properties they exhibit and how people respond to them.