A Pilot Study on the Effects of Oblique Strategies on Creativity and Divergent Thinking using an Intervaled Alternative Uses Task

Wahagn Mkrtchyan
Leiden University
Leiden, Noord-Holland, Netherlands
w.k.mkrtchyan@umail.leidenuniv.nl

ABSTRACT

Oblique Strategies are prompt cards, created by Brian Eno and Peter Schmidt, designed to remove creative blocks by encouraging lateral thinking. While the cards are well known internationally and by the scientific community, there is a lack of scientific investigation on their degree of effectiveness and use in multiple domains. To address this knowledge gap, an analysis was first performed on the Oblique Strategies to characterize the characteristics of the prompts on the cards, and assess their applicability to a large set of domains. Using this categorization of the prompts, three of the original prompts were chosen for a pilot study. To study the effects of these prompts an altered version of the Alternative Uses Task was developed, which splits the task across two intervals with a short break in-between. To study the effectiveness of the prompts two variations of the test are used, one where a prompt is shown during the break and one where no prompt is shown. The timing of the break was chosen such that the participants are assumed to have mostly run out of ideas. The study aims to determine if showing a prompt during the break assists participants in producing more and/or qualitatively better answers in the second interval after the break. Measures of the number of answers and their originality, flexibility and elaboration scores are used to evaluate the quality of the response. Moreover, the possibility of computer generating an endless amount of Oblique Strategy-like cards, possibly for a specific creative stage or domain is considered. Based on the original Oblique Strategies, new prompts are generated using a state-of-the-art transformer model. Three generated strategies similar to the three original Oblique Strategies were chosen based on their characteristics. The generated prompts are evaluated next to the original strategies in the same Alternative Uses Task. Despite not being able to provide a conclusive answer on the merits of Oblique Strategies in a divergent thinking task it is shown that in this trial test the generated prompts did not under perform compared to their original counterparts. Additionally, through this initial pilot study, it is shown that participants generally perceived the small break during their answering period as useful in allowing them to think of better or more answers after the break (presumably by removing fixation) this was also (significantly) reflected in their performance scores after the break.

KEYWORDS

Creativity, Divergent Thinking, Oblique Strategies Effectiveness, Creative Prompt Card Generation, Alternative Uses Task with Intervals

1 INTRODUCTION

Oblique Strategies (OS) are prompt cards created in the 1970s by musician Brian Eno and artist Peter Schmidt. The cards have been used by artists and musicians alike, and the first edition from 1975 is a true collector's item [8]. Physically, it takes the form of a deck of printed cards (as shown in figure 1) in a black case. Each card offers a challenging constraint intended to help artists break creative blocks by encouraging lateral thinking.

Don't be frightened of cliches

Figure 1: Oblique Strategy Card

The cards were most famously used by Eno during the recording of David Bowie's Berlin Trilogy albums. They were used again on Bowie's 1995 album Outside, which Eno was involved with as a writer, producer and musician [14]. Carlos Alomar, who worked with Eno and Bowie on all these albums, was a fan of using the cards, later saying 'at the Center for Performing Arts at the Stevens Institute of Technology, where I teach, on the wall are Brian Eno's Oblique Strategies cards. And when my students get a mental block, I immediately direct them to that wall.' [14]. This illustrates the popularity of the prompt cards.

Moreover, in the decades since the OS cards were first released, new decks intended to encourage creativity have hit the market [4]. Notable releases were Roger von Oech's Creative Whack Pack in 1989, Naomi Epel's The Observation Deck in 1998, and Matt Vojacek's Game of Creativity in 2018. Among this set of decks were also IDEO's Method Cards released in 2003 that focus on ways to ensure people remain the center of the design process.

While the creators of the OS cards may have had a variety of aims when producing the cards (e.g., generating economic value), facilitating the creative process was amongst their primary goals. Nevertheless, scientific investigation into the effective use of these

cards is noticeable by its absence. Consequently, the impact of OS cards on the performance of creative tasks is an open scientific question.

A large amount of research is focused on the categorization of creative processes. The problem of studying creative processes has been approached along multiple dimensions including mental attitudes, stages and types of creative thought. The stages of creative thought considered by a majority of today's scientific research are preparation, incubation, intimation, illumination and verification [13]. Moreover, in a recent paper by Glăveanu [5] a framework is presented which attempts to set a new standard for classifying different creative processes with a new set of categories for creative expression. Consequently, of any method attempting to facilitate a creative process it should first be determined which stages and types of creative processes it addresses.

To establish the stages addressed by OS, a characterization table was prepared with the original set of cards to categorize the characteristics of the evocative phrases on the prompts. The goal of this exercise was to determine the stage(s) of creative thought (preparation, incubation, intimation, illumination, verification) a phrase relates to and which type of creative expression(s) (actor, action, artifact, audience, affordances) it concerns. Furthermore the nature of the phrases was categorized by denoting if the OS is vague or concrete and whether the prompt can be perceived as a call to action, a general statement and/or a question. Moreover, the applicability of the cards in a large set of domains was considered. A smaller subset of 10 cards was then evaluated with three extra judges to ensure that the initial review was objective to a sufficient degree. An approximately 50% agreement between the combined vote of external judges compared to the initial classifications was deemed to be sufficient to use the initial categorization, as a point of reference, for later stages.

Additionally, through generating strategies, by feeding a state-of-the-art transformer named InferKit[1] original strategies as input, three OS-like strategies were selected. After selecting this set of three generated strategies similar to the three chosen original Oblique Strategies, based on their characteristics, the generated prompts were evaluated next to the original strategies.

While upon initial empiric investigation the different inputs had only slight effect on the creative characteristics of the outcome, the produced lists by InferKit often included items that could be mistaken for one of the original OS because of their similar nature. A handpicked subset of elements from various generated lists were evaluated next to the original strategies in an Alternative Uses Task (AUT) to compare their performance. The AUT designed by J.P. Guilford in 1967 [6] is a widely used test where participants are asked to think of as many uses as possible for a simple object, for example a brick or a paperclip.

A pilot test was performed to benchmark the effects of the cards, both the original and computer generated ones, with two AUTs.

---

[1]https://inferkit.com/

To minimize bias the two chosen objects used for the AUTs (brick, paperclip) were shown to participants in the order first brick then paperclip or vice versa first paperclip then brick. The two possible orders were as equally distributed between participants as possible. Each object had two intervals of three and subsequently two minutes, during which the participants could list uses for the AUT, with a break in between those intervals of 30 sec.

In this second break for $\frac{1}{3}$rd of all participants no strategy was shown similar to the first break. For the remaining $\frac{2}{3}$rds a strategy was shown and suggested to be used after the break with half of the times presenting one of three predetermined original OS created by Eno and the other half presenting one of three predetermined similar strategies picked from the output of InferKit's generated lists. The prompt was shown in the break at the point in time where the test's participant was assumed to have mostly run out of ideas. The goal being to determine if showing this phrase in the break could help the participant form more answers in the second interval after the break and whether those answers would be more original, flexible and/or elaborate. No significant increase or decrease in performance was measured for the participants who received an original OS nor for the combination of participants who received either an OS or OS-like prompt, compared to the users who only received a break with no prompt attached.

The remainder of the paper is organized as follows. In section 2, the related literature used in this study is reviewed. The data on the OS used and the AUTs is presented in section 3 and 4 respectively while the goal and the testing approach is presented in section 5. In section 6 the methods of evaluation are defined and in section 7 the experimental results are presented. A discussion on further work is shown in section 8 and the paper is concluded in section 9.

2 RELATED WORK

While this study revolves around Oblique Strategies, the small quantity of research into this topic were not of significant relevance. Despite the limited research regarding OS however, a fast amount of scientific articles have been written on categorizing creative processes as well as Alternative Uses Tasks and factors influencing it's outcome.

2.1 Characterization

To have a scientific understanding of the creativity aspects of the original set of prompt cards created by Eno these cards were examined and categorized as the first step in this study. The characteristics upon which the cards were identified were decided upon using empirical research and the popular categorization methods for creative processes.

The creative process has been studied as consisting of various dimensions of thought such as mental attitudes, stages of creative thought, and types of creative thought as early as 1985 [3]. To characterise the OS cards, Sadler-Smith's five stage model was used [13], which expands Wallas's original four stage model [15] (preparation, incubation, illumination and verification) by arguing for

an additional 'intimation' stage between the 'incubation' and 'illumination' stages. According to Sadler-Smith this stage is a partly conscious stage between 'incubation' and 'illumination' as Wallas originally intended [15]. Therefore the prompts were categorized in stages of creative thought of Preparation, Incubation, Intimation, Illumination and Verification.

Furthermore the strategies were identified upon by a recent adaptation by Glăveanu on the four P's model [12] by Rhodes named the five A's [5]. Glăveanu argues that for multiple decades the psychology of creativity has been based on the four P's framework (person, process, product and press) however, this framework has limitations. The limitations Glăveanu presents are the discreteness of the elements in the four P's framework. The connection between each of the four P's is not considered in the original framework. He claims that interrelations need to be made explicit. Also Glăveanu's aim is to expand the context, between creativity and societal and cultural elements to explicitly include material or physical 'press'.

Glăveanu describes the 5 A's framework as an expansion of the 4 P's framework as follows: Person becomes Actor, Process becomes Action, Product becomes Artifact and Press is split into Audiences and Affordances. Significantly, the actor exists only in relation to an audience, action cannot take place outside of interactions with a social and material world, and artifacts embody the cultural traditions of different communities. This framework of Actor, Action, Artifact, Audience, Affordances was incorporated in this current study into our categorization of Eno's OS.

2.2 Alternative Uses Task

An AUT is usually time constrained. It is a divergent thinking test, which tests the participant's ability to think creatively using four different measures for evaluation of the uses provided by the user. These criteria are listed below.

- the number of proposed uses
- the number of unique categories or domains addressed
- the originality of each use
- the elaboration of the description of each use.

A noteworthy study on AUTs is a research on the computational treatment of the AUT answers that was developed by the authors of the research 'Computational Extraction of Metrics and Normative Data on the Alternative Uses Test on a Set of 420 Household Objects' [11]. Olteteanu et al. argue in this paper that the usual ad hoc performance ratings which most studies conducting an AUT use bring multiple disadvantages, besides a high manual work requirement.

Examples of pitfalls that is stated in this research are firstly comparing fluency scores on answers of different objects, because one object could have intrinsically more creative and specifically fluency potential then another object. The second mentioned pitfall is that the originality score might not reflect the true originality of the answers. This can be the case when in a very creative group several users present a unique use that would be considered more unoriginal as when presented in a less creative group.

The links provided in this research to the computational program which would calculate creativity scores based on mathematical clustering were not valid and could therefore not be utilized.

Additionally, one study conducted an AUT experiment to determine the effects of the visual design of the response box on creative divergent thinking [10]. Mainly, Mohr et al. claim to have proven that decisions made about the visual design of a response box influence the number, length, and quality of the responses participants give in a creativity task. Mohr et al. found that variations in the size, number, and type of the response box, which is omitted in most literature has an influence on the fluency, elaboration, and originality of the output of participants engaged in the creative thinking task. In our current experiment a large unsegmented box was used. This is in line with the aforementioned paper [10] which is concluded by stating that when seeking a wide range of scores to differentiate respondents, their results suggest the best response box to use would be a large unsegmented box. This design would better allow for each participant to be expressive with their answers, in their own way.

3 OBLIQUE STRATEGIES DATA

The data used to produce the characterizations of the OS consisted of the original strategies created by Eno [7].

A data table was created (included in appendix section A.1) where each strategy was identified using a mix of criteria. The first part of which were selected through an initial study, which was conducted to establish baseline criteria to capture the differences between strategies. To locate the strategies in relation to established research frameworks the criteria in five A's framework [5] and creative stages model [13] were included. These were used to categorize which of five A's paradigms (Action, Artifact, Actor, Audience and Affordances) a given strategy is synonymous with and the creative processes (Preparation, Incubation, Intimation, Illumination, Verification) each card concerns. For each strategy the following questions were asked and answered with a binary yes or no per question.

The questions devised through an initial empirical study were the following. One question was on the generality of the card. It denotes whether the card is applicable (interpretable) in a large set of creative domains or only in one domain / setting / discipline. Also, three questions were presented where the given phrase is identified as being a call to action, question and/or a statement. Lastly, a question was posed asking whether the idea presented in the prompt is a concrete one, meaning that it's specific and not abstract and/or vague, which does of course leave room for personal interpretation.

A set of five questions were presented where it is determined if the card concerns any of the following: Actor (personal attributes in relation to a societal context), Action (coordinated psychological and behavioral manifestation), Artifact (cultural context of artifact production and evaluation), Audience (the interdependence

between creators and the social world) and Affordances (the interdependence between creators and the material word).

Then another set of five questions were presented where the card is identified by which stages it could concern, namely the creative process stages stated by Sadler-Smith citing Wallas' intentions: Preparation, Incubation, Intimation, Illumination and Verification.

The preparation and verification stages are conscious mentally regulated staged. Logic, mathematics, experiments and observations are typical aspects of preparation and verification. Preparation moreover, relates also to gathering information strategies while verification is a way to evaluate the creative work.

Incubation can have one of two attributes. The first is not to voluntarily or consciously think on a particular problem. The second attribute is that a series of unconscious and involuntary mental events may take place. As for the first attribute, abstention from mental work may take one of two forms, conscious mental work on other problems (distraction) and relaxation from all mental work.

Intimation is a semi-conscious process where the thought forms naturally before the idea is as complete as possible and afterwards consciously the essence of this process is captured before it has drifted away. Illumination is more of a singular moment where the 'click' happens and everything comes together to present the solution.

The entire set of questions (in the order they were applied) is summarized below.

- Is the card applicable (interpretable) in a large set of creative domains or categories?
- Does the card related to an action, artifact, actor, audience, affordances?
- Which of the following creative processes is the card synonymous with: preparation, incubation, intimation, illumination, verification?
- What type of phrase is it, call to action, question, statement?
- Is the presented idea a concrete one (rather than abstract)?

3.1 Agreement Matrix

To investigate the degree of objectivity of the characterizations 10 of the cards were presented to three judges. The prompts were manually picked on the basis of varying categorizations attempting to include every possibility in terms of the nature of the phrase and which process or paradigm it's been identified to relate to, according to the initial classification to be evaluated. The 15 questions described above were asked to the judges which revolve around the creative domain, type of process and message aspects of the cards.

An agreement matrix (see appendix section A.2) was created based on answers submitted by the judges to identify if a common ground of how the cards are perceived exists. Overall, a roughly 50% agreement was found. However, on a per column basis, the

agreement was higher, this was more so the case on the unambiguous questions such as if the phrase is a question or not.

The same phrase was provided twice to identify and discard any submissions that would be randomly filled in. However, with a recalculation of the agreement of the 9 unique phrases a large shift in the scores was not observed and the overall agreement remained approximately 50%, with several small score differences on a per column basis.

The characterization was deemed objective enough to utilize in the later stages of the study as a point of reference for picking OS cards based on their characteristics.

3.2 OS-like Prompt Generation

The possibility of computer generating an endless amount of OS-like cards was interesting due to the possibilities computer generated prompts could offer such as catering for a specific creative domain or stage of thought. Based on the characterization table the text generation tool InferKit was utilized to generate new OS-like prompts. This tool uses a transformer model [9] and given a text snippet generates the subsequent words that it determines would follow the input text. The transformer is an architecture for transforming one text sequence into another one with the help of two parts: the encoder and decoder [9]. The encoder maps the input sequence to symbols it can understand and the decoder reads the symbols and expands on them with similar symbols which can be converted to human readable text afterwards. By providing it several lists of original OS in various amounts (3-10), ranging from similar cards according to their characterization to a more diverse set, the lists are completed at each iteration by the transformer. The produced items were a continuation of the provided list with list items that the model determined would follow.

4 ALTERNATIVE USES TASK DATA

Two AUTs were deployed to benchmark the effectiveness of prompts in a divergent thinking exercise involving a setup on Amazon's Mechanical Turk for recruitment and a Qualtrics[2] survey to run the experiment.

4.1 Mechanical Turk Setup

Participants were recruited via Amazon's Mechanical Turk [3] service. A total of 39 participants completed the study for 0.35$ USD compensation per participant. Recruitment was restricted to participants mostly located in Northern America, Oceania and Western Europe (AU, AT, BE, CA, DK, FI, FR, DE, IE, LU, NL, NO, SE, CH, GB, US) with a Human Intelligence Task (HIT) approval rate (%) greater than 98 and number of HITs approved greater than 50. The age range was not restricted.

4.2 Participants

Of the 39 completed responses many were declined due to invalid responses which mostly involved copying a list of alternative uses

---

[2]https://www.qualtrics.com/
[3]https://www.mturk.com/

found on the internet[4], while several others were removed for not providing any answer on one or more of the four intervals (answering periods) before and after the break on either of the two objects (brick, paperclip). Several times a user had provided answers in all intervals, however the provided answers after the break were essentially duplicates of answers previously presented before the break for the same object. Because duplicate answers were removed during the processing of the responses (the second occurrence was removed), these participant's responses also had blank responses and were therefore left out of calculations. While technically possible to completely run out of ideas before the break, due to the experiment's design, it is highly unlikely for any participants engaged in the task to not be able to provide a single valid additional answer in the second interval after the break. If this happens it is a sign of poor engagement in the task and therefore the entirety of the response was not considered. After removing all the invalid responses 16 participants remained.

4.3 Responses

The collected responses were tabulated with a single row for each participant. The row first contained the information on the participant (e.g, unique MTurk id, gender, age) then the answers for the four intervals, two per object (before and after the break) for both objects and the order in which the objects appeared. The provided responses were manually processed and divided in respective elements of single answers. Also, per participant the outcomes of the multiple choice questions were noted, regarding the user's perception of the break and/or strategy, which were asked at the end of the experiment. All duplicate uses were removed, either from the same interval or duplicates in regard to the answers before and after the break for the same object (e.g., paperclip) for the same participant. If the answer was similar but not exactly the same it was not deleted. Moreover, if as a result of this operation the response for the given interval remained empty the submission of the given participant in its entirety was removed from the data table and not processed further. Also, the specific prompt, which was presented to the participant, was noted. This data table was called the 'Responses' data table.

4.4 Project Data

The entirety of data used in this study, including for the purposes of further research, can be found on GitHub[5].

5 PROBLEM DEFINITION

The aim of this pilot study was to measure the effect an OS or OS-like prompt can have on divergent thinking through the use of a divergent thinking exercise, in the form of an AUT.

The experiment in the form of a survey (see appendix section A.4) contained two AUTs and ended with questions regarding the participants' experience regarding the experiment. The AUTs had two notable differences. The first difference was the object that the

users had to list uses for, one of the AUTs concerned the object brick while the other concerned the object paperclip. The second difference was that while both AUTs were divided in two intervals with a break in between the break for the second AUT contained in $\frac{1}{3}$rd of the cases a prompt with an OS and in another $\frac{1}{3}$rd of the cases an OS-like prompt that the users were suggested to use after the break.

5.1 Selected Prompt Cards

There were six possible prompts that were shown during the break in the second AUT task. Three of the prompts were original OS created by Eno while the remaining three were computer generated using a transformer model. Apart from the original strategies by Eno the generated prompts were used to compare the performance of the two types of strategies to each other.

A set of three original strategy's were chosen (see figure 2) based on maximizing variance in their characterization in the 'characteristics' data table devised in the first stage of this study. The aim being to include different types of strategies to discover if a particular type of strategy affects divergent thinking differently. After the selection of the original strategies a set of three generated prompts (using InferKit), similar to the the chosen original OS based on their characteristics, were selected to be evaluated next to the original strategies in the AUTs (see figure 3). The characterizations of the OS and OS-like prompts can be found in appendix section A.3.

Figure 2: Original Oblique Strategies

- Twist the spine
- What wouldn't you do?
- Don't be frightened of cliches

Figure 3: Computer Generated OS-like strategies

- Draw a door. It should open to 3
- Do not worry about what others will say
- Change your routine, your fears

5.2 AUT Experiment Design

In the survey after a short introduction text stating the purpose of the experiment and insisting to avoid using external resources when answering, participants were randomly assigned to one of three experimental conditions that determined the strategy presented in the second break (no strategy, Oblique Strategy or generated OS-like strategy). All participants were asked to specify their gender and age first.

Additionally, the order of the two objects (brick then paperclip afterwards or vice versa paper clip then brick afterwards) seen in the divergent thinking task was equally and randomly distributed across participants. Participants were given 5 min to complete the task (with a short break appearing after 3 minutes for a duration of

---

[4]https://www.enviromate.co.uk/blog/twenty-incredibly-creative-ways-reuse-old-bricks, https://www.bobvila.com/slideshow/12-ways-you-never-thought-to-use-a-paper-clip-51028

[5]https://github.com/Wahagn/Oblique-Strategies-Effectiveness

30s) and after completing the AUT for both objects were automatically advanced to a set of questions regarding their perception of the breaks. Additionally their experience regarding the presented strategy (if one was presented to them) was informed about. This setup is shown in figure 4 below.

The multiple choice questions included asking the participant if they had run out of ideas before the break occurred, and if in their perception the break was useful in coming up with new and/or better answers after the break. Moreover, they were asked, in the case that an OS or OS-like strategy was presented to them, how much they used the strategy and in case they had not used it why not. Lastly, all participants that had received a strategy were asked to judge, in the case they used the strategy at least a little, how much if at all it improved in their perception the fluency (amount), flexibility (concerning multiple domains or categories), originality (uniqueness) and elaboration (amount of detail) of their answers after the break the strategy was presented in.

The pitfalls mentioned in the work of Olteteanu et al. [11] described in the related work section 2 were addressed in our experiment. Bias in fluency scores between different objects and originality scores between groups of creative and less creative groups was avoided. This was achieved first of all by providing all participants with the same two objects in a roughly equal distribution on the order that the two objects came in. Furthermore, direct creativity (including fluency) was calculated, however only the differences in creativity before and after a small break for the same object for the same participant were evaluated. Also, the amount of responses and the general amount of answers per response to be evaluated was small enough for the manual labor to be manageable.

The study was conducted online using the Qualtrics survey tool. The design of the experiment can be found in appendix section A.4 the survey as a whole including it's branching flow and further details can be found in the GitHub repository, as presented in section 4.4.

6 METHODS OF EVALUATION

6.1 Matching

The responses from all participants from the 'Responses' table were split in two separate tables for the answers on the two objects brick and paperclip respectively, stating in each table the unique uses presented by the participants for the given object.

Uses that were similar in function were grouped together. For the item 'brick' this meant that stating you can build a building with it is one type of use. Regardless of the fact that one person might state 'building a school' and another 'building a restaurant' both would match with the same row e.g, 'buildings (tower / church / school / restaurant / store / lighthouse)'. For the object 'paperclip' this meant that skewering food, whether it's a kebab or a marshmallow is grouped together. Also the uses 'carving', 'whittling' and 'chisel' are essentially so similar in function that they were grouped together. So are the uses 'clipping paper', 'organizing paper' and 'holding paper', however not the use 'staple' as this specification

while similar in usage and function is different by being loaded in a stapler and it might be used to hammer into a surface and to hold something in place, on multiple types of surfaces however, such as paper, wood or a cork board. You can find the entire matching table in the appendix in section A.5.

This grouping together ensured that when calculating the originality of an answer we don't award high scores for a use that is almost identical as another specified use, by grouping them together and noting the occurrence of both uses in one item resulting in a single, higher and thus worse occurrence value. The way that the occurrence was specified transpired by traversing the list of values of the users in the 'Responses' data table. The position of the occurrence was noted behind each grouped use in this 'Matching' data table according to the response box. If the item was specified as the first answer a '0;' was noted and if it was the second answer a '1;' was noted etc. If the grouped use occurred another time in the same response or in another person's response, again the new position was noted behind the already noted occurrences e.g, adding a 3 to the '0;' it became '0;3;' etc. . This meant that the use had occurrence two as it occurred two times, in positions 0 and 3 respectively. In the paper by Olteteanu et al. [11] these positions were used in the originality calculation by differentiation on answers that were creative, meaning only having few occurrences and mostly being presented as one of user's last answers, and being specified by the user as one of their first answers in contrast, as a sign of a creative mind. Despite not using these position lists in the same manner in our current experiment, they were used to identify and match the original responses to this matching data table for score calculation.

Furthermore, of each item (grouped use) in the current data table it was determined which of the five following domains it belonged to. The categories were Tool, Accessory, Construction Material, Philosophical, Miscellaneous. Each group of similar uses was attached to one (and only one) of these categories and these domains were determined through observed differences in the respondents' presented answers and existing practices for AUT task evaluation noted in the paper by Olteteanu et al. [11].

6.2 Scoring

Scores for the performance of respondents in stating alternative uses for the two objects were awarded using the regular practices for scoring used for an AUT. The evaluation of answers of an AUT is generally calculated using the following metrics:

Fluency - Fluency is the number of uses the participant can present.

Flexibility - Flexibility is the number of conceptual domains the answers relate to. In this case Tool, Accessory, Construction Material, Philosophical, Miscellaneous are utilized as categories answers can relate to. The possible domains chosen for the flexibility were based on standards of AUT evaluation [11] and an empirical study of the provided answers.

ID	age	gender	Oblique strategy (assigned)	Oblique Item 3min. (1)	Break 30 sec.	Oblique Item 2min. (1)	Oblique Item 3min. (2)	Break 30 sec.	Oblique Item 2min. (2)
1	16	m	none	Brick	no strategy	Brick	Paperclip	assigned strategy presented	Paperclip
2	23	f	none	Paperclip	no strategy	Paperclip	Brick	assigned strategy presented	Brick
3	73	f	none	Brick	no strategy	Brick	Paperclip	assigned strategy presented	Paperclip
4	45	m	original	Brick	no strategy	Brick	Paperclip	assigned strategy presented	Paperclip
5	67	m	original	Paperclip	no strategy	Paperclip	Brick	assigned strategy presented	Brick
6	30	m	original	Paperclip	no strategy	Paperclip	Brick	assigned strategy presented	Brick
7	32	m	generated	Brick	no strategy	Brick	Paperclip	assigned strategy presented	Paperclip
8	41	f	generated	Paperclip	no strategy	Paperclip	Brick	assigned strategy presented	Brick
9	29	m	generated	Brick	no strategy	Brick	Paperclip	assigned strategy presented	Paperclip

Figure 4: Experiment Scheme

This rating was calculated by ticking each category which one of the answers related to e.g, if there are three answers presented and two concern a type of tool and one a form of jewellery two points would be awarded 1 for the domain tool and one for the domain accessory.

Originality - Originality measures how uncommon a stated use by a participant is compared to other participants. Counting the list of occurrences in the 'Matching' data table denoted how many times a certain use was stated. To attach a reversed relation between the awarded points and the occurrence count, number one was divided by the occurrence count. Novelty however, was not taken into account by employing human judges who would rate how novel they consider a particular use to be.

Elaboration - Elaboration is a measure of the detail that the answers posses. To calculate the elaboration score a simple strategy was used where the amount of words minus the stop words in English were counted. Of course the removal of the stop words was an extra feature to better match the intricate judging a human judge could perform in rating detail not only based on the sheer amount of words. Because a long answer does not necessarily equal a detailed answer.

6.3 Calculation of Scores

A script (see appendix section A.6) was created to calculate the scores. The script traverses all answers provided by the respondents in the 'Responses' data table and matches each provided use with one of the items in the 'Matching' data tables. Once the match is made points can be awarded to the respondent for their answer. After all uses respondents provided are matched the calculated points can be saved in a 'Calculations' data table.

The script has five stages namely parsing responses and creating a list of them, finding max values for fluency and elaboration for both the paperclip and brick responses separately, parsing the items in the grouped uses 'Matching' data tables and creating a list with them, calculating the results and saving the calculations.

6.3.1 Stage 1. In the first stage all responses from the participants are compiled in a list ordered per participant then per answering interval (two for both objects brick and paperclip, before and after the break). The list is saved for later use.

6.3.2 Stage 2. In the second stage the responses are traversed once again with a different goal this time. In this instance the maximum value for fluency and elaboration is determined for both objects separately. This is useful for later to normalize the scoring by dividing a score with the maximum points that any user obtained for the given object. In this process per participant per answering interval the amount of answers are counted and if the amount is higher than that of previous respondents the maximum value is updated for the respective object brick or paperclip. The same happens in regard to the elaboration score where we find the response with highest amount of words for the respective object. Provided items in the responses are stripped of stop words using the 'nltk' package's English stop words [1].

6.3.3 Stage 3. In the third stage a list is compiled of the 'Matching' data tables for the two objects. Every element in the list contains the data in the data table such as a representation of similar uses that occurred, as well as the occurrences per grouping of uses in the form of a position list denoting which *n*-th answer this use appeared as in the response boxes of the participants. This position list is later used to aid the matching process. Also saved per item is the conceptual domain the group of uses belongs to e.g, Tool, Accessory etc.

6.3.4 Stage 4. In the fourth stage the scores are calculated and awarded. Each element from the aforementioned list of answers formed in the first stage is matched against an element in the matching list from the third stage. The matching is done through a similarity scores using the Minimum Edit Distance (MED) [2] that denotes a measure in difference between two text sequences. This is possible because the items in the 'Matching' data tables represent a combination of very similar answers in the 'Responses' data table through generalisation. While they are representations, their text similarity should theoretically be closer than any of the other items in the 'Matching' tables.

Nevertheless, certain tactics are employed to aid this matching process such as double checking if the position the answer appears in corresponds to the last element of the occurrence positions noted in the position list of this item in the 'Matching' list. If so, this position is then deleted so the next time we match with this item we look at the next position this answer appeared as to check in our list of answers if this indeed corresponds with the position this answer appears in.

Furthermore, answers are matched up to the previously furthest matched item (and the next one) from the 'Matching' list. This is the case because it limits our choices to only valid items to match with. These 'Matching' data tables were created by subsequently traversing the 'Responses' data table from top to bottom. This means that when considering an answer of a participant it was only added as a new item in the 'Matching' data tables if it could not be grouped together with any of the previously noted groups of similar uses. So when our last matched item in the 'Matching' list had index 23 the next answer could either be one of these 22 items (when counting from 0), in case a very similar use appeared before. It could potentially also have been original enough to be added as the next item at index 24 (23 + 1) but no index higher than that would be valid.

However, even these aids were not sufficient to ensure a faultless matching procedure and due time constraint a strategy was chosen to hard code and overrule the few mismatches that still happened. A better solution would be to change the wording in the 'Matching' data tables to better represent the various provided uses it should match against to ensure the highest similarity scores for each corresponding item, preventing any mismatches.

Scores were awarded once the match was established. First of all the originality score for a single listed answer was calculated by equation 1.

$$\frac{1}{occurrence} \tag{1}$$

Occurrence is the total amount of times this (or almost identical) answers appeared. Subsequently, for every domain that was covered by the answers in a given response box a point was awarded. For example, if there were two answers and one is an accessory and the other is a specification of tool then two points would be awarded. A maximum of four points were awarded as none of the participants had an answer which was identified as being philosophical or a paradigm. Afterwards the fluency score was determined by counting the amount of answers in the response box while the elaboration score was calculated by adding up the amount of words in the response box and dividing it by the fluency score e.g., 53 words, 8 answers so the elaboration score is shown below in formula 2.

$$\frac{53}{8} \tag{2}$$

.

To normalize the results, again a division was performed on the four metrics of fluency, flexibility, originality and elaboration. All four of these scores were divided by the highest amount of the point for this metric, for the the item brick or paperclip, that a participant obtained. The flexibility score was merely a sum of 0.25 points for each of the five domains the answers in the response can address. This was the case because the highest score anyone obtained is 4 out of 5 domains meaning $0.25 \times 4 = 1$. For the elaboration and fluency scores we used the maximum values by any participant, found in the second stage of this calculation script. These values were used to divide the scores as shown in the equations 3 and 4

below, with maximums for the respective objects brick or paperclip the response related to.

$$fluency\_score\_normalized = \frac{fluency\_score}{max\_fluency\_score} \tag{3}$$

$$elaboration\_score\_normalized = \frac{elaboration\_score}{max\_elaboration\_score} \tag{4}$$

The originality score was averaged over all the originality scores (see formula 1) per provided answer in the interval, using the fluency score to divide by. A perfect score would be of course if all answers were unique, a full point would be awarded.

Afterwards an average score was calculated of the normalized scores for fluency, originality, flexibility and elaboration. The normalized scores all have a range between 0 and 1 and are rounded to two decimals before calculating the average. Also, the average itself is rounded to two decimals after its calculation.

6.3.5 Stage 5. Lastly the calculated and rounded scores (including the averages) are saved in a new data table and later manually transferred to the original 'Responses' data table to be used for participant performance analysis. The scores are structured in 20 items, (4 (metric measures) + 1 (average)) × 4 = 20 items meaning all four metrics (fluency, flexibility etc.) for each answering interval (two for both objects brick and paperclip item, before and after the break).

For the python calculation script see appendix section A.6.

6.4 Performance Change Calculations

In the 'Responses' data table extra columns were added to calculate the delta of the performance before and after the break on each object (brick, paperclip) for all respondents. This metric denotes by what percentage the performance increased or decreased after the break compared to the performance before the break with equation 5.

$$delta\_break = \frac{avg\_perf\_after\_br - avg\_perf\_before\_br}{avg\_perf\_before\_br} \times 100 \tag{5}$$

. This metric was calculated for the averages of the two AUTs of brick and paperclip (delta_break1+2) as well as for the two AUTs separately (delta_break1) and (delta_break2) respectively, regardless of the assigned object in the first or second AUT (paperclip or brick). The delta_break2 measure was chosen as the most interesting value given that the second AUT included the break where the strategy was shown in the form of an OS or OS-like prompt, to $\frac{2}{3}$rds of participants. Additionally, a second and third measure for this break were calculated using, instead of the delta difference, a fraction difference as shown in equation 6 and a differential as follows in equation 7 , again using the averages from the calculation script, to correctly asses the performance difference before and after the

break.

$$fraction\_break = \frac{avg\_perf\_after\_br}{avg\_perf\_before\_br} \qquad (6)$$

$$differential\_break = avg\_perf\_after\_br - avg\_perf\_before\_br \qquad (7)$$

The data table was loaded in JASP and several tests were performed. First off all descriptive statistics were generated for the study population regarding their age, gender and how the strategies were assigned. Furthermore, it was checked if the order of the AUTs, brick then paperclip or paperclip then brick, had any dependence on the assigned strategy. Afterwards the delta_break2 was subjected to an ANOVA test to determine if type of assigned strategy (OS or OS-like) had an effect on the performance increase or decrease after the break. As a second check the same ANOVA test was run on the differential_break2 this time to affirm the result of the previous test. Afterwards, a T-test was run to compare the delta_break2 performance between participants who had received a strategy, regardless of the type (original or computer generated) to participants who had only received a break with no strategy. A last ANOVA test was conducted to determine if the perception of respondents regarding the usefulness of the breaks corresponded to their performance increase or decrease after the breaks.

7 RESULTS

7.1 Distribution Specifics

7.1.1 Gender. The results of the gender frequency table 1 below show that our experiment was completed mostly by females, 70%, compared to 30% males. Other gender options such as non-binary or 'prefer not to say' were not chosen.

Table 1: Frequencies for gender

gender	Frequency	Percent	Valid Percent	Cumul. Percent
F	11	68.750	68.750	68.750
M	5	31.250	31.250	100.000
Missing	0	0.000		
Total	16	100.000		

7.1.2 Age. The results of the a age statistics table 2 below show that our experiment was completed by respondents between ages 23 and 62 with a mean of 35.

Table 2: Descriptive Statistics age

	age
Mean	35.688
Std. Deviation	11.170
Minimum	23.000
Maximum	62.000

7.1.3 Strategy vs No Strategy. The results of the 'a_strategy' frequency table 3 below show that a large percentage, 80%, of our participants received a strategy in the second break compared to the respondents who received no strategy in either of the two breaks.

Table 3: Frequencies for a_strategy

a_strategy	Frequency	Percent	Valid Percent	Cumul. Percent
0	3	18.750	18.750	18.750
1	13	81.250	81.250	100.000
Missing	0	0.000		
Total	16	100.000		

7.1.4 Strategy. The results of the strategy type frequency table 4 below show that the variance between the types of strategies that the participants received is lower than the variance between all strategies and no strategies in the previous table 3. The three original OS strategies grouped together as strategy 1 and the three computer generated OS-like strategies denoted by strategy number 2 had a distribution of 30% to 50% respectively.

Table 4: Frequencies for strategy

strategy	Frequency	Percent	Valid Percent	Cumul. Percent
0	3	18.750	18.750	18.750
1	5	31.250	31.250	50.000
2	8	50.000	50.000	100.000
Missing	0	0.000		
Total	16	100.000		

7.2 Order vs Strategy

The result of the chi-squared test below in table 6 assured us through a large p-value that no significant dependency existed between the order of the AUTs brick then paperclip or paperclip then brick and the assigned strategy for the second break. This test was used to rule out that the distribution was skewed towards one of the objects for cases with a strategy in the second break. Thus the object could have no significant effect on the performance values. It was expected from the perfect 8 vs 8 distribution of the order and the slight variations in assigned strategies 0, 1 and 2 for each order in table 5.

Table 5: Contingency Tables order vs strategy

	strategy			
order	0	1	2	Total
0	1	2	5	8
1	2	3	3	8
Total	3	5	8	16

Table 6: Chi-Squared Tests

	Value	df	p
X^2	1.033	2	0.597
N	16		

7.3 Strategy Types Compared

An ANOVA test was performed both on the measures delta_break2 and additionally differential_break2 which relate to the performance increase or decrease after the second break (in the second AUT object) and the type of strategy that the respondents received. The types of strategies included the three possibilities of not receiving a strategy (strategy 0 in the table), receiving an original OS (strategy 1 in the table) and receiving a computer generated OS-like strategy (strategy 2 in the table). As shown in the results of the first and second ANOVA test regarding the delta_break2 measure and differential_break2 measure in table 7 and table 8 respectively no significant relation was found.

This means that we can't prove that the type of strategy had effect on the performance increase or decrease after the second break. The means for the differential_break2 in table 10 suggest the OS to be less useful than having a computer generated strategy or no strategy at all, the means for the delta_break2 in table 9 suggest the same. The case 0 denoting no strategy shows an increase in performance of 9% while the case of receiving an original OS shows a slight decrease of 1%. Furthermore, the mean regarding strategy 2, a computer generated prompt, shows an increase in performance after the break of 12%, suggesting that the computer generated strategies were more useful than having no strategy or an original OS created by Eno.

Table 7: ANOVA - delta_break2

Cases	Sum of Squares	df	Mean Square	F	p
strategy	564.052	2	282.026	0.363	0.702
Residuals	10095.843	13	776.603		

Table 8: ANOVA - differential_break2

Cases	Sum of Squares	df	Mean Square	F	p
strategy	0.005	2	0.002	0.156	0.857
Residuals	0.205	13	0.016		

7.4 Strategy vs No Strategy

Through a T-test in table 11 on the 'a_strategy' column denoting if the participant received a strategy or not it became apparent that also no effect on the performance could significantly be related to the increase or decrease of the performance of participants on the second AUT. Thus a significant positive effect of the prompts, both OS and OS-like was not measured, regardless of the measure used

Table 9: Descriptives - delta_break2

strategy	Mean	SD	N
0	9.395	14.790	3
1	-0.952	37.888	5
2	12.458	23.653	8

Table 10: Descriptives - differential_break2

strategy	Mean	SD	N
0	0.033	0.057	3
1	-0.006	0.186	5
2	0.031	0.093	8

such as the delta, differential and fraction of performance averages before and after the second break. Nor a significant negative effect was noted as the p-values was too large, however, the trend seemed to be negative with a mean of 9% performance increase after the break for the group that only received a break compared to the 7% increase in performance for the group that received one of the two types of strategies as shown for delta_break2 in table 12.

Table 11: Independent Samples T-Test strategy vs no strategy

	t	df	p
delta_break2	0.119	14	0.907
frac_break2	0.119	14	0.907
differential_break2	0.210	14	0.837

Table 12: T-Test Group Descriptives

	Group	N	Mean	SD	SE
delta_break2	0	3	9.395	14.790	8.539
	1	13	7.300	29.171	8.091
frac_break2	0	3	1.094	0.148	0.085
	1	13	1.073	0.292	0.081
differential_break2	0	3	0.033	0.057	0.033
	1	13	0.017	0.130	0.036

7.5 Participant's Perspective

An ANOVA test was also performed to determine if the respondents answers who indicated that the breaks were useful corresponded to their performance scores after the break. The multiple question was phrased in the following way:

Having a small break in between writing answers helped me come up with new/better ideas after the break.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree

- Somewhat disagree
- Strongly disagree

The ANOVA test requires a minimum of two examples for each possible case and the data lacked in this regard as not all answers were chosen twice or more in total. Therefore, the first two possible answers were grouped together as a positive response (denoted by a 1 in the table) and the last three answers were grouped together as a negative answer (denoted by a 0 in the table). Moreover, because the question related to the break in general the average performance delta measure of both breaks together was used, namely delta_break1+2.

The ANOVA test confirmed that the respondents who perceived the breaks as useful had on average a better performance (10%) after the break as shown in table 13. Moreover, the group that did not perceive the break as useful performed worse after the break (-4%). With a p-value of 0.05 as seen in table 14 this effect could be considered significant. Noteworthy is also the fact that more participants belong to the group that found the break to be useful.

Table 13: Descriptives - delta_break1+2

break_positive	Mean	SD	N
0	-3.517	11.836	7
1	9.531	12.203	9

Table 14: ANOVA - delta_break1+2

Cases	Sum of Squares	df	Mean Square	F	p
break_positive	670.446	1	670.446	4.619	0.050
Residuals	2031.885	14	145.135		

8 DISCUSSION

The results of our small sample study suggest that offering a break in between answering periods increases the quantity and quality performance scores of answers when allowing the participant of a divergent thinking task to continuing devising answers after the break. Moreover, the results show that participants in general could correctly identify whether the break had a positive effect on the quantity and/or quality of their answers after the break.

Nevertheless, our pilot study had its limitations. The most important limitation was the fact that despite employing an experiment design with a between-subjects (original vs computer generated strategies vs no strategy) as well as within-subjects (break with strategy vs break without strategy) approach it lacked the usual minimum set of 30 participants. Moreover, the scoring methods involved manual judgement from one judge only, decreasing the objectivity of the calculated scores. In addition, the small amount of respondents did not allow for a in-between phrases calculation to score the phrases on effect.

For further research it would be useful to address the aforementioned limitations. First of all the experiment could be run with a larger population size, for example with 500 participants. A large sample size would help to prevent bias in the population in terms of their creativity, allowing the group to better represent a national or world average. This would make the measured scores of fluency, flexibility, originality and elaboration more objective given that we are normalization each score with the best score achieved in the group.

Furthermore, the scoring could be further improved by using a mathematical clustering approach as suggested in the paper by Olteteanu et al. [11] to prevent subjectivity in the choice of flexibility domains and remove the ad hoc clustering of similar uses. If this approach would not be feasible for a next study a first step of the introduction of (more) human judges would also increase confidence in rating elaboration and originality of answers.

This research would additionally benefit by making use of a large amount of participants that phrases could be scored individually or in groups of similar phrases. If positive effect(s) or negative effect(s) could be proven to exist regrading the use of OS or OS-like prompts in a divergent thinking task, it would allow to narrow it down and asses which (types of) phrases have a positive/negative effect.

Lastly, as the two AUTs employed in this experiment were intervaled with a break in between, there was no effect to be measured between an answering period with and without a break, of the same length. In this study it was measured how well the participants performed after the break in comparison to before the break. The incorporation of measuring effects the addition of the break has could prove or disprove the benefits a period of 'incubation' or 'illumination' [13] could have for a divergent thinking task.

9 CONCLUSION

We have found no scientifically significant evidence that presenting an Oblique Strategy in a break of an intervaled divergent thinking task can improve the results after the break, over only providing a break with no strategy. This might suggest that the period of incubation that the break presumably provides where users do not voluntarily or consciously think on a particular problem was as, if not more, useful as the same break with a prompt attached.

Moreover, it is apparent from the numbers in our ANOVA tests regarding effects of the types of strategies that the computers generated prompt could not be proven to be less or more effective than their original counterparts.

Nevertheless, we have no evidence to conclude, due to the nature of the survey's setup, that a break itself aids in improving results overall. This stems from the fact that all participants were provided with the break and this break was the same length of time for all participants on both the first and second AUT items. No assumption can be made on the potential performance scores achieved if no break had occurred at all.

When looking at the perception of the participants however, we can conclude that most participants in this pilot study perceived the break to be useful in aiding them to devise more or better answers after the break. The participants who indicated the break to be useful indeed increased their performance scores after the break.

REFERENCES

[1] 2021. Removing stop words with NLTK in Python - GeeksforGeeks. https://www.geeksforgeeks.org/removing-stop-words-nltk-python [Online; accessed 18. Aug. 2021].

[2] T. Bex. 2020. FuzzyWuzzy: Fuzzy String Matching in Python | Towards Data Science. *Medium* (Nov 2020). https://towardsdatascience.com/fuzzywuzzy-fuzzy-string-matching-in-python-beginners-guide-9adc0edf4b35

[3] Michael Cagle. 1985. A General Abstract-Concrete Model of Creative Thinking. *The Journal of Creative Behavior* 19, 2 (1985), 104–109. https://doi.org/10.1002/j.2162-6057.1985.tb00643.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2162-6057.1985.tb00643.x

[4] Artsy Editorial and Abigail Cain. 2018. Dealing with Creative Block? A Deck of Cards Might Help. https://www.artsy.net/article/artsy-editorial-dealing-creative-block-deck-cards-help

[5] Vlad Petre Glăveanu. 2013. Rewriting the Language of Creativity: The Five A's Framework. *Review of General Psychology* 17 (03 2013), 69. https://doi.org/10.1037/a0029528

[6] Joy Paul Guilford. 1968. The Nature of Human Intelligence. *Science* 162, 3857 (1968), 990–991. https://doi.org/10.1126/science.162.3857.990-a arXiv:https://science.sciencemag.org/content/162/3857/990.2.full.pdf

[7] Joel Parker Henderson. 2018. joelparkerhenderson/oblique-strategies. https://github.com/joelparkerhenderson/oblique-strategies

[8] Colin Marshall. 2018. Behold the Original Deck of Oblique Strategies Cards, Handwritten by Brian Eno Himself. https://www.openculture.com/2018/12/behold-original-deck-oblique-strategies-cards-handwritten-brian-eno.html

[9] Maxime. 2020. What is a Transformer? - Inside Machine learning - Medium. *Medium* (Sep 2020). https://medium.com/inside-machine-learning/what-is-a-transformer-d07dd1fbec04

[10] Alicia Hofelich Mohr, Andrew Sell, and Thomas Lindsay. 2016. Thinking Inside the Box: Visual Design of the Response Box Affects Creative Divergent Thinking in an Online Survey. *Social Science Computer Review* 34, 3 (2016), 347–359. https://doi.org/10.1177/0894439315588736 arXiv:https://doi.org/10.1177/0894439315588736

[11] Ana-Maria Olteteanu, Richard Hass, and Faheem Hassan Zunjani. 2019. Computational Extraction of Metrics and Normative Data on the Alternative Uses Test on a Set of 420 Household Objects. *Behavior Research Methods* (07 2019).

[12] Mel Rhodes. 1961. An Analysis of Creativity. *The Phi Delta Kappan* 42, 7 (1961), 305–310. http://www.jstor.org/stable/20342603

[13] Eugene Sadler-Smith. 2015. Wallas' Four-Stage Model of the Creative Process: More Than Meets the Eye? *Creativity Research Journal* 27, 4 (2015), 342–352. https://doi.org/10.1080/10400419.2015.1087277 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/10400419.2015.1087277

[14] Jeff Slate. 2016. Carlos Alomar's Golden Years with David Bowie (Interview). https://web.archive.org/web/20160213015838/http://www.rockcellarmagazine.com/2016/02/08/david-bowie-carlos-alomar-interview-golden-years/

[15] Graham Wallas. 1926. *Art of Thought*. Harcourt Brace, New York, NY.

A   APPENDIX
A.1   Oblique Strategies Characterization Table

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
Abandon normal instruments	yes	yes	no	no	no	yes	yes	no	no	no	no	yes	no	no	no
Accept advice	yes	yes	no	no	yes	no	yes	no	no	no	no	yes	no	no	no
Accretion	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
A line has two sides	yes	no	no	no	no	yes	no	yes	no	no	no	yes	no	no	no
Allow an easement (an easement is the abandonment of a stricture)	yes	yes	no	no	no	no	yes	no	yes	no	no	yes	no	no	no
Are there sections? Consider transitions	yes	yes	no	no	no	no	no	no	no	yes	no	yes	no	no	no
Ask people to work against their better judgement	yes	yes	no	yes	no	no	yes	no	no	no	no	yes	no	no	no
Ask your body	yes	yes	no	yes	no	yes	no	no	yes	no	no	yes	no	no	no
Assemble some of the instruments in a group and treat the group	no	yes	no	yes	yes	no	yes	no	no	yes	no	yes	no	no	no
Balance the consistency principle with the inconsistency principle	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Be dirty	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Breathe more deeply	yes	yes	no	yes	no	no	no	yes	no	no	no	yes	no	no	yes
Bridges -build -burn	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Cascades	yes	no	no	no	no	yes	no	yes	no	yes	no	no	no	yes	no
Change instrument roles	no	yes	no	no	no	yes	yes	no	no	yes	no	yes	no	no	no
Change nothing and continue with immaculate consistency	yes	yes	no	no	no	no	no	yes	no	no	no	yes	no	no	no
Children's voices -speaking -singing	no	no	yes	no	no	no	no	no	yes	no	no	yes	no	no	no
Cluster analysis	yes	yes	no	no	no	no	yes	no	no	no	no	no	no	yes	no
Consider different fading systems	no	yes	no	no	no	no	yes	no	no	no	no	yes	no	no	no
Consult other sources -promising -unpromising	yes	yes	no	no	yes	no	yes	no	no	yes	no	yes	no	no	no
Convert a melodic element into a rhythmic element	no	yes	yes	no	no	no	no	no	no	yes	no	yes	no	no	no
Courage!	yes	yes	no	yes	no	no	no	no	yes	no	no	yes	no	no	no
Cut a vital connection	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Decorate, decorate	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no	yes
Define an area as `safe' and use it as an anchor	yes	yes	no	no	no	yes	yes	no	no	yes	no	yes	no	no	no
Destroy -nothing -the most important thing	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
Discard an axiom	yes	yes	no	no	no	no	yes	no	no	no	no	yes	no	no	no
Disconnect from desire	yes	yes	no	no	no	no	no	yes	yes	no	no	yes	no	no	no
Discover the recipes you are using and abandon them	yes	yes	no	no	no	yes	yes	no	no	no	no	yes	no	no	no
Distorting time	yes	yes	no	no	no	no	no	no	no	no	no	no	no	yes	no
Do nothing for as long as possible	yes	yes	no	no	no	no	no	yes	no	no	no	yes	no	no	yes
Don't be afraid of things because they're easy to do	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Don't be frightened of cliches	yes	yes	no	no	no	no	yes	no	no	no	no	yes	no	no	no
Don't be frightened to display your talents	yes	yes	no	yes	no	no	no	no	no	no	no	yes	no	no	no
Don't break the silence	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Don't stress one thing more than another	yes	yes	yes	no	no	no	no	no	no	yes	no	yes	no	no	no
Do something boring	yes	yes	no	no	no	no	no	yes	no	no	no	yes	no	no	yes
Do the washing up	yes	yes	no	no	no	no	no	yes	no	yes	no	yes	no	no	no
Do the words need changing?	yes	no	yes	no	no	no	no	no	no	no	yes	yes	yes	no	yes
Do we need holes?	yes	no	yes	no	no	no	no	no	no	no	yes	no	yes	no	no
Emphasize differences	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Emphasize repetitions	yes	yes	yes	no	no	no	no	no	yes	no	no	yes	no	no	no
Emphasize the flaws	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Faced with a choice, do both (given by Dieter Rot)	yes	yes	no	no	no	no	no	no	no	yes	no	yes	no	no	no
Feedback recordings into an acoustic situation	no	no	yes	no	no	yes	no	no	no	yes	no	yes	no	no	no
Fill every beat with something	yes	yes	no	no	no	yes	no	no	no	yes	no	yes	no	no	no
Get your neck massaged	yes	no	no	yes	no	no	no	no	no	no	no	yes	no	no	no
Ghost echoes	no	no	yes	no	no	yes	no	no	no	yes	no	no	no	yes	no
Give the game away	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Give way to your worst impulse	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Go slowly all the way round the outside	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Honor thy error as a hidden intention	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
How would you have done it?	yes	yes	no	yes	no	no	no	no	yes	no	no	no	yes	no	no
Humanize something free of error	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
Imagine the music as a moving chain or caterpillar	no	yes	no	yes	no	no	no	no	no	no	no	yes	no	no	no
Imagine the music as a set of disconnected events	no	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Infinitesimal gradations	yes	yes	no	no	no	no	yes	no	no	no	no	no	no	yes	no
Intentions -credibility of -nobility of -humility of	yes	yes	no	no	no	no	no	no	no	no	no	no	no	yes	no
Into the impossible	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Is it finished?	yes	no	yes	no	no	no	no	no	no	no	yes	no	yes	no	yes
Is there something missing?	yes	no	yes	no	no	no	no	no	yes	no	yes	no	yes	no	no
Is the tuning appropriate?	yes	no	yes	no	no	no	no	no	no	no	yes	no	yes	no	no
Just carry on	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Left channel, right channel, centre channel	yes	no	no	no	no	yes	no	no	no	yes	no	yes	no	no	no
Listen in total darkness, or in a very large room, very quietly	no	yes	no	no	no	yes	no	no	no	no	no	yes	no	no	no
Listen to the quiet voice	yes	yes	no	no	no	yes	no	no	yes	no	no	yes	no	no	no
Look at a very small object, look at its centre	yes	yes	no	no	no	no	no	yes	no	no	no	yes	no	no	no
Look at the order in which you do things	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	yes
Look closely at the most embarrassing details and amplify them	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	yes
Lowest common denominator check -single beat -single note -single riff	no	yes	no	no	no	yes	no	no	no	no	no	yes	no	no	no
Make a blank valuable by putting it in an exquisite frame	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Make an exhaustive list of everything you might do and do the last thing on the list	yes	yes	no	no	no	no	yes	no	no	no	no	yes	no	no	yes
Make a sudden, destructive unpredictable action; incorporate	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Mechanicalize something idiosyncratic	yes	yes	yes	no	no	yes	no	no	yes	no	no	yes	no	no	no
Mute and continue	no	yes	no	no	no	no	no	no	no	no	no	yes	no	no	no
Only one element of each kind	yes	yes	yes	no	no	yes	yes	no	yes	yes	no	yes	no	no	no
(Organic) machinery	yes	yes	no	no	no	yes	no	yes	yes	no	no	no	no	yes	no

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
Overtly resist change	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	yes
Put in earplugs	no	yes	no	yes	no	no	no	no	no	yes	no	yes	no	no	yes
Remember those quiet evenings	yes	yes	no	no	no	no	no	yes	no	no	no	yes	no	no	no
Remove ambiguities and convert to specifics	yes	yes	no	no	no	yes	no	no	yes	no	no	yes	no	no	no
Remove specifics and convert to ambiguities	yes	yes	no	no	no	yes	no	no	yes	no	no	yes	no	no	no
Repetition is a form of change	yes	yes	no	no	no	no	no	no	yes	no	no	no	no	yes	no
Reverse	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Short circuit (example: a man eating peas with the idea that they will improve his virility shovels them straight into his lap)	yes	yes	no	no	no	no	yes	no	no	yes	no	yes	no	no	no
Shut the door and listen from outside	no	yes	no	no	no	yes	no	no	no	yes	no	yes	no	no	yes
Simple subtraction	yes	yes	no	no	no	no	no	no	yes	no	no	no	no	yes	no
Spectrum analysis	yes	yes	no	no	no	no	yes	no	no	no	no	no	no	yes	no
Take a break	yes	yes	no	yes	no	no	no	yes	no	no	no	yes	no	no	yes
Take away the elements in order of apparent non-importance	yes	yes	no	no	no	no	no	no	no	yes	yes	yes	no	no	no
Tape your mouth (given by Ritva Saarikko)	no	yes	no	yes	no	yes	no	no	no	yes	no	yes	no	no	yes
The inconsistency principle	yes	yes	no	no	no		yes	no	yes	no	no	no	no	yes	no
The tape is now the music	no	yes	no	no	no	yes	no	no	yes	no	no	yes	no	no	no
Think of the radio	yes	yes	no	no	no	yes	no	yes	yes	no	no	yes	no	no	no
Tidy up	yes	yes	no	no	no	no	no	yes	yes	yes	no	yes	no	no	no
Trust in the you of now	yes	yes	no	yes	no	no	no	no	no	no	no	yes	no	no	no
Turn it upside down	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
Twist the spine	yes	yes	no	no	no	yes	no	no	yes	no	no	yes	no	no	no
Use an old idea	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no	yes
Use an unacceptable color	yes	no	yes	no	no	yes	no	no	no	yes	no	yes	no	no	no
Use fewer notes	no	no	yes	no	no	yes	no	no	no	yes	no	yes	no	no	yes
Use filters	no	no	yes	no	no	yes	no	no	no	yes	no	yes	no	no	yes
Use "unqualified" people	yes	no	no	yes	no	no	no	no	no	yes	no	yes	no	no	yes
Water	yes	no	no	no	no	yes	no	yes	no	no	no	no	no	yes	no
What are you really thinking about just now? Incorporate	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
What is the reality of the situation?	yes	yes	no	no	no	no	yes	no	yes	no	yes	no	yes	no	no
What mistakes did you make last time?	yes	yes	no	yes	no	no	yes	no	no	no	no	no	yes	no	no
What would your closest friend do?	yes	yes	no	yes	yes	no	no	no	yes	no	no	no	yes	no	no
What wouldn't you do?	yes	yes	no	yes	no	no	yes	no	yes	no	no	no	yes	no	no
Work at a different speed	yes	yes	no	yes	no	no	no	no	no	yes	no	yes	no	no	yes
You are an engineer	yes	yes	no	yes	no	no	no	no	yes	no	no	no	no	yes	no
You can only make one dot at a time	yes	yes	no	yes	no	no	no	no	yes	no	no	yes	no	no	no
You don't have to be ashamed of using your own ideas	yes	yes	no	yes	no	no	no	no	no	no	no	yes	no	no	no
[blank white card]	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	yes	yes	yes	yes

Characteristics of Oblique Strategies:	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	statement	question	concrete	FIELD17	agreement % per row
Cluster analysis	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	no (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	yes (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)		agreement: 50.0%
Make a sudden, destructive unpredictable action; incorporate	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:100.0%, n:0.0%)	no (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	no (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	no (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)		agreement: 45.0%
What wouldn't you do?	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	yes (y:33.3%, n:66.7%)	no (y:33.3%, n:66.7%)	no (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)		agreement: 50.0%
Use ""unqualified"" people	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	yes (y:66.7%, n:33.3%)	no (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	no (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)	yes (y:33.3%, n:66.7%)	no	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	yes (y:66.7%, n:33.3%)		agreement: 55.0%
Twist the spine	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	yes (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)		agreement: 45.0%
Assemble some of the instruments in a group and treat the group	no (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	yes (y:0.0%, n:100.0%)	yes (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	no (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	yes (y:100.0%, n:0.0%)		agreement: 50.0%
Don't be frightened of cliches	yes (y:100.0%, n:0.0%)	yes (y:33.3%, n:66.7%)	no (y:33.3%, n:66.7%)	no (y:66.7%, n:33.3%)	no (y:100.0%, n:0.0%)	no (y:33.3%, n:66.7%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	no (y:33.3%, n:66.7%)	no (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	no (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)		agreement: 45.0%
Change nothing and continue with immaculate consistency	yes (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)	yes (y:33.3%, n:66.7%)	no (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:100.0%, n:0.0%)	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:100.0%, n:0.0%)		agreement: 55.0%
Left channel, right channel, centre channel	yes (y:33.3%, n:66.7%)	yes (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)	yes (y:66.7%, n:33.3%)	no (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	yes (y:66.7%, n:33.3%)	yes (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	yes (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)		agreement: 55.0%
Twist the spine	yes (y:66.7%, n:33.3%)	yes (y:100.0%, n:0.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)	no (y:66.7%, n:33.3%)	no (y:33.3%, n:66.7%)	yes (y:33.3%, n:66.7%)	no (y:0.0%, n:100.0%)	no (y:33.3%, n:66.7%)	yes (y:100.0%, n:0.0%)	no (y:0.0%, n:100.0%)	no (y:0.0%, n:100.0%)	no (y:66.7%, n:33.3%)		agreement: 45.0%
agreement % per column	agreement: 80.0%	agreement: 80.0%	agreement: 50.0%	agreement: 50.0%	agreement: 40.0%	agreement: 50.0%	agreement: 60.0%	agreement: 60.0%	agreement: 50.0%	agreement: 70.0%	agreement: 40.0%	agreement: 100.0%	agreement: 100.0%	agreement: 100.0%	agreement: 60.0%		total agreement: 49.5%

A.3 OS vs. OS-like prompts

Characteristics:	original	general	action	artifact	actor	audience	affordances	preparation	incubation	intimation	illumination	verification	call to action	question	statement	concrete
Twist the spine	yes	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no	no
What wouldn't you do?	yes	yes	yes	no	yes	no	no	yes	no	yes	no	no	no	no	yes	no
Don't be frightened of cliches	yes	yes	yes	no	no	no	no	yes	no	no	no	no	yes	yes	no	no
Draw a door. It should open to 3	no	yes	yes	no	no	no	no	yes	no	yes	no	no	yes	no	no	no
Do not worry about what others will say	no	yes	yes	no	no	yes	no	yes	no	no	no	no	no	no	yes	no
Change your routine, your fears	no	yes	yes	no	yes	no	no	yes	no	yes	no	no	yes	no	no	no

A.4 Survey Design

Introduction

Dear participant,

This is a survey for a thesis project regarding computers and creativity. Our goal is to study the effect of a method for encouraging creative thinking. The survey has two simple and straightforward tasks and will in total take approx. 15 min to complete. At the last stage, you will get various multiple-choice questions on your experience doing the exercise. **PLEASE COMPLETE THE SURVEY ALONE WITHOUT THE USE OF ANY EXTERNAL RESOURCES, OTHERWISE, YOUR SUBMISSION WILL BE DECLINED.** There are no right or wrong answers, your experience is the most important factor.

Personal Details

Please select your gender

Male
Female
Non-binary / third gender
Prefer not to say

Please select your age

| | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| age |

Explanation (1)

In the next part you will be asked to name as many possible uses as you can for an ordinary object. You'll have 5 minutes to do so with a 30 sec. break in between after 3 minutes. When you are ready to start go to the next page.

Brick (1)

0 3 0 0

Name as many uses of a Brick as possible

Paperclip (1)

0 3 0 0

Name as many uses of a Paperclip as possible

Break (1)

Have a little break! Please wait until the timer goes off.

3 0

Brick (2)

0 2 0 0

Please continue! Keep naming as many other uses of a Brick as possible.

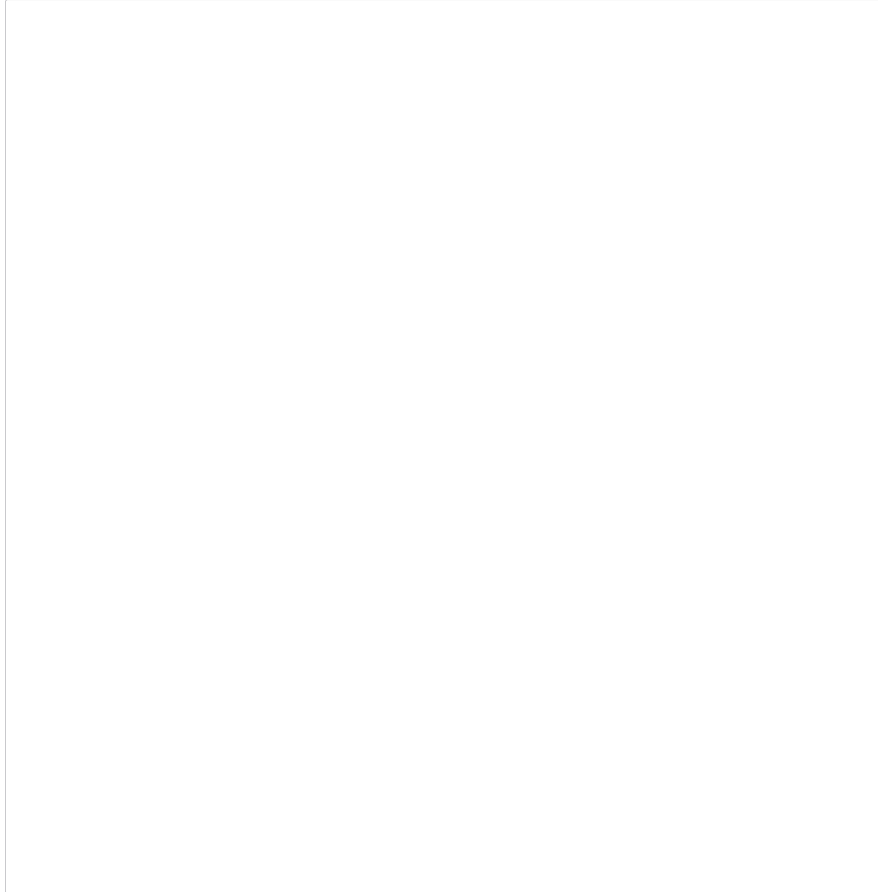**Explanation (2)**

In the next part you will again be asked to name as many possible uses as you can for a different ordinary object this time. You'll have 5 minutes to do so with a 30 sec. break in between after 3 minutes. When you are ready to start go to the next page.

Break (2) No strategy

Have a little break! Please wait until the timer goes off.

30

Break (2) Original strategy 1

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**Twist the spine**"

3 0

Break (2) Original strategy 2

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**What wouldn't you do?**"

3 0

Break (2) Original strategy 3

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**Don't be frightened of cliches**"

3 0

Break (2) CG strategy 1

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**Draw a door. It should open to 3**"

3 0

Break (2) CG strategy 2

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**Do not worry about what others will say**"

3 0

Break (2) CG strategy 3

Have a little break! Please wait until the timer goes off.

After the break, if you run out of ideas, consider using the following phrase as a strategy for coming up with new ideas. You can interpret the phrase as you wish.

"**Change your routine, your fears**"
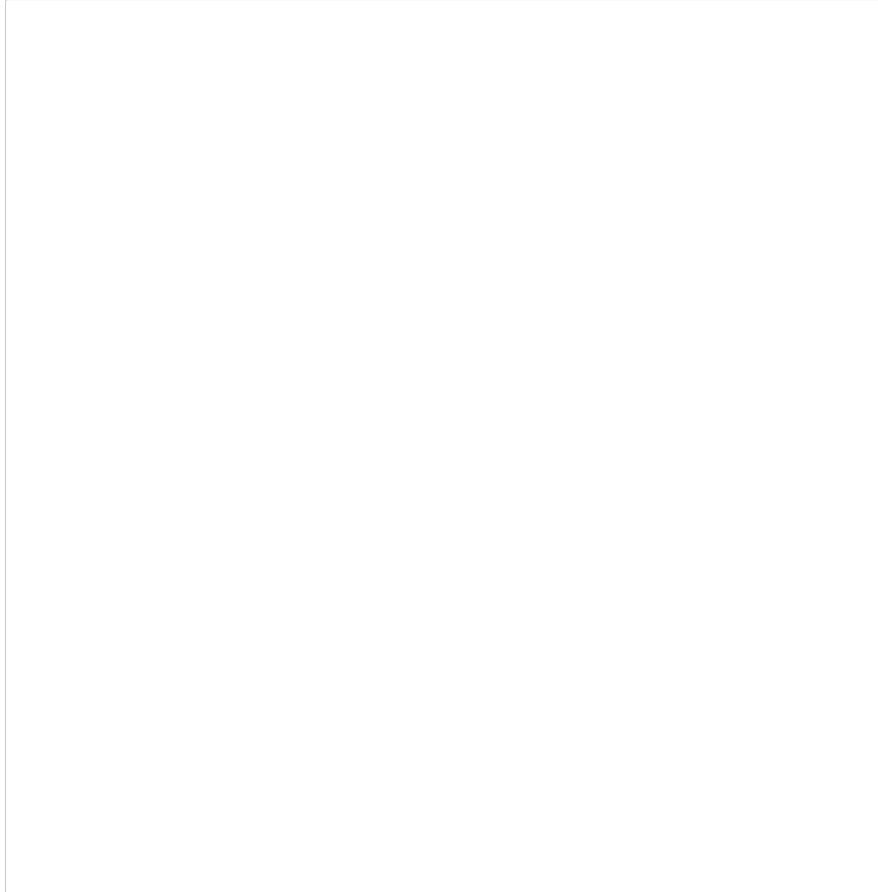
3 0

Paperclip (2)

0 2 0 0

Please continue! Keep naming as many other uses of a Paperclip as possible.

Multiple choice questions (use of break)

When I was presented with a break the second time, I had mostly run out of ideas.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

Having a small break in between writing answers helped me come up with new/better
ideas after the break.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

Multiple choice questions (strategy)

I have used the strategy that was presented during the second break.

A great deal

A lot

A moderate amount

A little

None at all

What is the reason you haven't used the strategy? (multiple answers possible)

I couldn't interpret the phrase in any meaningful way

I needed more time to interpret and use the strategy

Every time I ran out of ideas, a new idea popped up in my mind before I could use the strategy

Other reason, write below

The strategy that was presented during the second break helped me think of more ideas after the break.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

The strategy that was presented during the second break helped me come up with more original or unusual ideas after the break.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

The strategy that was presented during the second break helped me come up with ideas in a broader range of categories and applications after the break.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

The strategy that was presented during the second break helped me come up with ideas with a higher level of detail after the break.

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

MTurk Code

Here is your id number (survey code for mturk):

${e://Field/Random%20ID}

Please copy this number. Once you've copied your ID, click on to the next page to finish the survey. After that you can insert your id number on MTurk as your survey code.

Powered by Qualtrics

A.5 Matching data table

A.5.1 Brick.

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
build house / home construction	0;1;0;0;0;0;1;0;1;1;0			x		
chimney	1;2			x		
build wall	2;0;1;1;2;4			x		
floor construction	3;1			x		
paper weight	4;0;2;13;0;1;0	x				
(melee) weapon	5;1;0;1	x				
projectile	6					x
build deck	7			x		
build flowerpot	8			x		
build barbecue	9			x		
brick (garden) path / walkway	0;5;3;3			x		
planter holder	1			x		
garden bench	2;5			x		
waterfall construction	3;6			x		
birdbath	4;7			x		
yard art	5			x		
brick fire pit	6;3;3;5;3;0			x		
use (fitness) weights (lifting)	2;6;10;5;1					x
pavement construction	3;3			x		
throwing	0;4					x
candle holder	1	x				
flatten something (dough)	2;7	x				
book end	2;0;2;1	x				
coconut opener	3	x				
door wedge / doorstop	4;0;0;4;3	x				
hammer	5;12	x				
car pedal holder	7	x				
build oven	1			x		
insect killer	2	x				
skin exfoliator	3	x				
can opener	4	x				
bottle opener	5	x				
balloon popper	6	x				
to build / construction / brickwork	0;2;3;0;1			x		
planter / plant holder / flowerpot	1;1;5;1;3;4			x		
to break a window	3;2	x				
hold down (light) objects (firewood tarp) fly away	4;0;7;0	x				
trace for (sidewalk) drawing	5;12;4	x				
pen / pencil / marker / toothbrush / lipstick / candle / beer holder	6;3;4;5;14;0;0	x				
to paint on (surface)	8;11					x
for stacking (them)	9					x

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
hit person / someone	1;1	x				
(walkway) decoration	2;3					x
reed diffuser stick holder	3	x				
fill holes with jello	4	x				
as a pet (put a string on it)	5					x
to block	1	x				
to attack	2	x				
to organize	0					x
to obstruct	1	x				
to restrict	3					x
to damage / scratch	4;14	x				
buildings (tower / church / school / restaurant / store / lighthouse)	2;6;7;8;9;0			x		
build porch	4			x		
pick it up	15					x
hit it like a drum	16	x				
sit on it / chair / stool / outdoor seat	17;10;1;5					x
keep car (from rolling) / stop car	1;3	x				
tealight holder	2;2	x				
(spaghetti) measure	4	x				
(car) booster seat	5	x				
put in toilet tank	4					x
picture frame	2					x
drying rack	6	x				
stepping stool (be taller)	7;2;4;2					x
balloon holder	8	x				
phone holder	9	x				
foot rest	11					x
pillow	13					x
make ladder	0			x		
sanding weight	2	x				
car jack	3	x				
DIY projects	1			x		
building a garden	2			x		
building furniture	4;1			x		
(indoor) fireplace	1;2;2			x		
build a door	0			x		
adding weight to objects	5;10	x				
sinking objects	6	x				
break / smash objects	7;6	x				
blocking a view	8	x				
negative ballast	1	x				

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
basemen (baseball)	3	x				
cooking (BBQ) stand	0;0			x		
outdoor fence	3			x		
nut cracker	3	x				
podium	2			x		
yoga block	5					x
(mini) sled for dolls	8		x			
mice shelter	9			x		
garden border	0			x		
edging a walkway	2			x		
crush and use for artwork	4					x
(theater) play prop	6					x
hit something	0	x				
bird nest	2			x		
weight machine test tool	6	x				
delimit a zone	7	x				
fill a hole	8					x
hide thing	1					x
play with	2					x
elevate hot water tank	1	x				

A.5.2 Paperclip.

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
key ring	0;8		x			
(tiny) screwdriver	1	x				
bra strap hider	2	x				
bookmark	3;6;7;1;3;0	x				
brush hair remover	4;3;4	x				
bag opener	5	x				
cherry pit remover	6	x				
smartphone stand	7			x		
lock pick	8;1;0;3;3;4;2;6;1	x				
antenna	0			x		
electrical circuit	1			x		
compass (needle)	2			x		
cork board tack	3	x				
tattoo needle	4;9	x				
(tiny) doll accessories (coat hanger / sword)	5;6			x		
fingertip spikey	7		x			
weapon	8;6	x				
(punk-rock) collar / hat spikes	9;10		x			
fishing hook	11			x		
mobile hanger	12			x		
build mini mobile	13			x		
pry out SD card / dvd drive	14;2	x				
stress relief toy (fidget)	15;7					x
tiny kite frame for clay model	16			x		
hanging objects (picture frames)	17;9;3;13;2			x		
paper airplane / frisbee weights	18;19					x
(hold) hair (bobby) clip	20;4;2;0;1;2;0;0;1	x				
keep / hold / organize / clip paper together	21;0;0;3;0;0;0;0;0;0;0;3	x				
bracelet	22;2;5;2		x			
slingshot ammo	23					x
rubber band shooter	24			x		
making music (shake box of paperclips)	25;10	x				
(tiny) piano keys	26			x		
(string) hooks on guitar	27			x		
sim card tool	0;1;12	x				
toothpick (remove stuff from teeth)	1;14	x				
(handcuff) key	1	x				
fishing lure	2	x				
make a ring	3;3		x			
poke holes (paper / potato / play-doh)	4;1;1;3;3	x				
splint for injured finger	5					x

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
eat food with (utensil)	6	x				
scratch (itchy) skin (back / toe)	7;10;13;8	x				
pop a balloon	1	x				
use to stitch clothing together	2	x				
trade for something	3					x
sell (on eBay)	4					x
phone (port) / tiny gun cleaner	0;1;4	x				
reset / restart modem / phone	2;4;1	x				
make earring	5;6;2;6;4;1		x			
making necklace	6;4;3;5		x			
hanging ornament	8;1;3	x				
magnet tester	11	x				
building a (stick) figures (animals / characters)	12;5			x		
(painting) decoration	14;9;10					x
throwing (if frustrated)	0;9					x
tracing (oval) designs (on paper) / drawing aid	2;4;4	x				
open (long line) for hard to reach places (fell behind desk)	3;4;3	x				
holding	1	x				
attaching	2	x				
to separate	0	x				
to stack	1					x
make a chain	3;6		x			
hold / pin (torn) clothes	5;2;2	x				
clip for reminders	0	x				
hold money / receipts together	1;2;0;2	x				
digging tool (for sand)	3;2;13	x				
(coffee) stirring tool	4;13	x				
fashion (straight) piece of metal (stick)	5;0			x		
cleaning brush	0	x				
bridge a blown fuse with	0			x		
replace zipper pull	1;9;1	x				
mark a page	4;5;4	x				
crochet marker	2	x				
nail	3	x				
pen holder	8	x				
crochet / knitting needles	10;11	x				
place marker	1	x				
chisel / whittling / carving tool	3;4;5	x				
painting tool (brush)	5;8	x				
wire / cord holder	6;0	x				
electricity experiment	7			x		

use	occurrence_positions	domain_tool	domain_accessory	domain_construction	domain_philosophical	domain_miscellaneous
repair items (glasses)	2;2					x
make nail art	4	x				
unclog (small) hole	1	x				
unlock door	1	x				
packet / bag closer	0;1	x				
clean nails (cuticle)	1;0;1;0	x				
poke someone	2	x				
making furniture	3	x				
connecting objects	1	x				
grabbing / gripping objects	2;12	x				
launching a mouse trap	5	x				
writing	6	x				
scratching (lottery card)	7;7	x				
hold / pin objects in place (flower)	11;2	x				
(roasting) skewered food (marshmallow / kebab)	2;7;8	x				
making sculptures	3			x		
making miniature houses	4			x		
flicking objects	5	x				
closing a (bracelet) clasp	4	x				
making figures for necklace (pendants and charms)	0			x		
conductor for metal (wires)	1;1					x
make curtain	3			x		
(bent) on a chair as practical joke	4					x
put gum on the end to collect items	7	x				
toy's teeth (alligators mouth)	1		x			
to roll around (a charger)	9					x
to untangle something	2	x				
to draw on sand	3	x				
staple	4					x
insect / bug killer	6	x				
nose ring	0		x			
self defense	5	x				
pin / thumbtack	7	x				
nappy clip	3	x				

A.6 Score Calculation Script

```python
#run with command 'python3 main.py'
#the script has to run without any mismatches/errors to succeed

import csv
from fuzzywuzzy import process
from nltk.corpus import stopwords


#-------- parse user inputs --------#

user_input = []
with open('results.csv') as csvfile:
    reader = csv.reader(csvfile, delimiter=',', quotechar='"')
    for row in reader:
        if(row[0] != ""):
            contents = []
            for col in range(5): #first 5 columns
                cell_value = None
                if(col > 0): #answer columns
                    cell_value = row[col].split(";")
                    cell_value.pop() #remove empty string at end
                else: #brickFirst col
                    cell_value = row[col]
                contents.append(cell_value)
            user_input.append(contents)

#-------- find max values --------#

fluency_score_max = [0,0]
originality_score_max = [0,0]
flexibility_score_max = [0,0]
elaboration_score_max = [0,0]

for row in user_input:
    for idx, cell_value in enumerate(row):
        if(idx > 0): #not brickFirst col
            fluency_score = [0,0]
            originality_score = [0,0]
            flexibility_score = [0,0]
            elaboration_score = [0,0]

            for use in cell_value:
                use = ' '.join([word for word in use.split()
                                if word not in (stopwords.words('english'))])
                if(idx < 3): #brick
                    fluency_score[0] = fluency_score[0] + 1
                    elaboration_score[0] = elaboration_score[0] + len(use.split())

                else: #paperclip
                    fluency_score[1] = fluency_score[1] + 1
                    elaboration_score[1] = elaboration_score[1] + len(use.split())

            for item in [0,1]: #adjust max scores both for brick and paperclip
```

```python
            if(fluency_score[item] >  fluency_score_max[item]):
                fluency_score_max[item] = fluency_score[item]

            if(fluency_score[item] > 0 and
                ((elaboration_score[item] / fluency_score[item]) >
                elaboration_score_max[item]) ):
                elaboration_score_max[item] = (elaboration_score[item] /
                                              fluency_score[item])


#-------- parse choices, pos list --------#

#general
choices = []
pos_list = []
occurence_list = []
domain_list = []

# frist process brick then paperclip exact same procedure
for filename in ['choices_brick.csv','choices_paperclip.csv']:
    with open(filename) as csvfile:
        reader = csv.reader(csvfile, delimiter=',', quotechar='"')
        choices_item = []
        pos_list_item = []
        occurence_list_item = []
        domain_list_item = []
        for row in reader:
            if(row[0] != "" and row[0] != "use"):
                choices_item.append(row[0])
                positions_item = row[1].split(";")
                positions_reverse_item = positions_item[::-1]
                pos_list_item.append(positions_reverse_item)
                occurence_list_item.append(len(positions_item))

                '''
                if(row[2] == "x"): domain_list_item.append(0)
                if(row[3] == "x"): domain_list_item.append(1)
                if(row[4] == "x"): domain_list_item.append(2)
                if(row[5] == "x"): domain_list_item.append(3)
                if(row[6] == "x"): domain_list_item.append(4)
                '''

                found_domain = False
                for d_i in range(2,7):
                    if(row[d_i] == "x"):
                        found_domain = True
                        domain_list_item.append(d_i - 2)

                if(not found_domain):
                    print("[ERROR] DID NOT FIND DOMAIN FOR ", row[0])
```

```
        #after all item choices processed
        choices.append(choices_item)
        pos_list.append(pos_list_item)
        occurence_list.append(occurence_list_item)
        domain_list.append(domain_list_item)

#-------- calculate scores --------#

calc_values = []

max_match_index = []
max_match_index.append(1) #for brick
max_match_index.append(1) #for paperclip

first_error = True

for user_id, row in enumerate(user_input):
    calc_row = ""
    b_p_switch = 0

    #if paperclip was first, switch the answers around so second item is brick
    if(row[0] == "0"):
        row[1:5] = row[1:5][::-1]
        row[1:3] = row[1:3][::-1]
        row[3:5] = row[3:5][::-1]
        b_p_switch = 1

    for idx, cell_value in enumerate(row):
        if idx > 0: # not the brick_first col
            if(idx == 3):#switching between lists for brick vs paperclip
                if(b_p_switch > 0): b_p_switch = 0
                else: b_p_switch = 1

            fluency_score = 0
            originality_score = 0
            flexibility_score = 0
            domains_score = [0,0,0,0,0]
            elaboration_score = 0

            for use_nr, use in enumerate(cell_value):
                use = '␣'.join([word for word in use.split()
                        if word not in (stopwords.words('english'))])
                matches = process.extract(use,
                        choices[b_p_switch][0:max_match_index[b_p_switch]],
                        limit=10)
                found = False
                matchIndex = 0

                for match in matches:
                    if not found:
                        matchIndex = choices[b_p_switch].index(match[0])
                        if(len(pos_list[b_p_switch][matchIndex]) and
```

```python
        pos_list[b_p_switch][matchIndex][-1] == str(use_nr)):
    found = True


    # overrule
    if(b_p_switch and  use == "using reach something"
        " fell behind desk open make one long line"):
        matchIndex = 57
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "build"):
        matchIndex = 33
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "Build lighthouse"):
        matchIndex = 52
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "Clip reminder front something"):
        matchIndex = 64
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "Stir something"):
        matchIndex = 67
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "Reset phone"):
        matchIndex = 48
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "pen holder"):
        matchIndex = 38
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "Build fire pit"):
        matchIndex = 16
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "construction"):
        matchIndex = 33
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "breaking objects"):
        matchIndex = 79
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "grabbing objects"):
        matchIndex = 91
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "poking objects"):
        matchIndex = 39
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "itching"):
        matchIndex = 42
        match = "MATCH_OVERRULED"
    if(not b_p_switch and use == "They used mini"
            " temporary stool (stacked top other)"):
        matchIndex = 56
        match = "MATCH_OVERRULED"
    if(b_p_switch and use == "Hold flowers place"):
        matchIndex = 95
        match = "MATCH_OVERRULED"
```

```python
                    if (b_p_switch and use == "hold_cords"):
                        matchIndex = 80
                        match = "MATCH_OVERRULED"


                    print(use, "->", match, "use_nr", use_nr, "pos",
                        pos_list[b_p_switch][matchIndex], "matchindex",
                        str(matchIndex) + "/"
                        + str(max_match_index[b_p_switch]-1) )
                    originality_score += 1/(occurence_list[b_p_switch][matchIndex])
                    domains_score[domain_list[b_p_switch][matchIndex]] = 1

                    if (matchIndex >= (max_match_index[b_p_switch] - 1)):
                        max_match_index[b_p_switch] = matchIndex + 2

                    pos_list[b_p_switch][matchIndex].pop()
                    if (not len(pos_list[b_p_switch][matchIndex])):
                        choices[b_p_switch][matchIndex] = "-----------"
                        "----------------------"
            if (use == "" or  (first_error and not found)):
                #print("[ERROR] DID NOT FIND FOR", use)
                print("[ERROR] DID_NOT_FIND_FOR", use, "IN", matches,
                "WITH", use_nr, "IN_POS", pos_list[b_p_switch][matchIndex])
                print(cell_value)
                print("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n")
                first_error = False


            fluency_score += 1
            elaboration_score += len(use.split()) #done after removing stopwords


    if (fluency_score > 0): #can't devide by 0 also 0 fluency = 0 originality, elaboration
        originality_score = round(originality_score / fluency_score, 2)
        elaboration_score = elaboration_score / fluency_score
    elaboration_score = round(elaboration_score / elaboration_score_max[b_p_switch], 2)
    fluency_score = round(fluency_score / fluency_score_max[b_p_switch] , 2)

    for domain in domains_score:
        if domain == 1: flexibility_score += 0.25
    avg = round ((
                    (fluency_score +
                    originality_score +
                    flexibility_score +
                    elaboration_score)/4
                ), 2)

    # print("User:\t", user_id, "(" + str(idx) +"/4)" )
    # print("Fluency:\t",        fluency_score)
    # print("Originality:\t", originality_score)
    # print("Flexibility:\t", flexibility_score)
    # print("Elaboration:\t", elaboration_score)
    # print("Total:\n", fluency_score, "," , originality_score , "," ,
    #flexibility_score , ",", elaboration_score)
```

```python
            calc_row += (str(fluency_score) +
                "," + str(originality_score) + "," + str(flexibility_score) +
                "," + str(elaboration_score) + "," + str(avg) + ",␣,")

        calc_values.append(calc_row + '\n')

#-------- save scores --------#

f = open("calculated.csv", "w")
for line in calc_values:
    f.write(line)
f.close()
```