Please cite this publication as follows:


Link to official URL (if available):

http://dx.doi.org/10.1016/j.tsc.2015.10.009

This version is made available in accordance with publishers’ policies. All material made available by CReaTE is protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.

Contact: create.library@canterbury.ac.uk
Beyond belief: Structured techniques prove more effective than a placebo intervention in a problem construction task

David Vernon*¹ and Ian Hocking

Affiliation

¹School of Psychology, Politics & Sociology, Canterbury Christ Church University, Canterbury, Kent. CT1 1QU, UK.

*Email: david.vernon@canterbury.ac.uk
Abstract

Problem construction is one of the first steps in creative problem solving and research has shown clear links between problem construction ability and creative output. Here, we compared two active techniques with that of a placebo intervention and show a benefit in problem construction performance for the active techniques. The active techniques required participants to either utilise six questions (six men), or adopt six perspectives, incorporating a range of specific questions (six hats). The placebo intervention (brain-breathing) was specifically constructed to seem both plausible and effective. We had 118 participants randomly allocated to one of the three groups (six men, six hats, brain-breathing) and, after reading a brief synopsis of their allocated tool, they then attempted to restate a given problem in as many different ways as they could within an allotted time. Performance was measured in terms of the fluency, quality, flexibility and originality of responses. Results showed that using the six men tool produced greater fluency, flexibility and originality relative to brain-breathing and the six hats. Use of the six hats tool also led to the production of more original responses relative to the brain-breathing control group. Importantly, there was no difference in reported motivation between the groups, but those using the six men and the brain-breathing tools found these easier to use compared to the six hats. Furthermore, those using the six men tool found this to be more useful and indicated that they were more likely to use this again in the future. Hence, both six men and six hats tools benefited performance, though in distinct ways. These results support the notion that explicitly scaffolding thinking can benefit creative problem solving.

Keywords: six good men, six thinking hats, placebo, creativity, problem construction.
Introduction

Creative problem solving (CPS) refers to a framework or approach used when attempting to solve a problem and produce both useful and original solutions. It is largely based on the early work of Osborn (1953), which stemmed from the desire to explicitly define the creative process and provide a structured approach to enhancing creative problem solving ability. Since then, whilst others have worked to develop and refine the framework (see e.g., Buijs, Smulders, & van der Meer, 2009; Isaksen & Treffinger, 2004; Puccio, Murdock, & Mance, 2005) there has been some general agreement that the process often begins with problem construction (see, Basadur, Graen, & Graen, 1982; Reiter-Palmon & Robinson, 2009). Problem construction includes the anticipation of problems, identifying problems when none are evident, and structuring an ill-defined problem so problem solving efforts can proceed (Mumford, Reiter-Palmon, & Redmond, 1994; Runco & Nemiro, 1994). Research has shown that problem construction is an essential skill in creative problem solving and that problem construction ability is clearly related to creative output (see, e.g., Mumford, Baughman, Threlfall, Supinski, & Costanza, 1996; Reiter-Palmon, Mumford, & Threlfall, 1998). Hence, attempts to train and/or improve problem construction ability would be expected to have beneficial effects on creative problem solving performance.

There is consensus in the literature that training in creative problem solving can be effective (see, DeHaan, 2009; Ma, 2006; McIntyre, Hite, & Rickard, 2003; Scott, Leritz, & Mumford, 2004b; Wang & Horng, 2002) and both business and education view such improvements as essential for future economic growth and educational development (see, Fontenot, 1993; Pithers & Soden, 2000). However, while there is much evidence that training programmes themselves can lead to improvements in problem finding, evidence for specific tools is patchy. Within the problem finding literature, we find evidence for brainstorming (e.g., Kurtzberg & Reale, 1999) and problem restatements (e.g., Mumford et al., 1994), as well as our own work on structured thinking techniques (Vernon & Hocking, 2014), but little else. This is surprising given the volume of tools that are out there, and the lack of a clear empirical foundation for such tools has recently led to calls for researchers to focus on this issue in an attempt to identify which tools work (see, Ma, 2006; Vernon, Hocking, & Tyler, under review).

This led us to examine whether training participants to use a specific tool would enhance their problem construction ability. The tools we focused on were, The Six Good Men
referred to by Rudyard Kipling (Kipling, 1902) and the Six Thinking Hats, put forward by de Bono (2009). The six men simply refer to the six questions: who, how, what, why, where and when. The rationale for this tool is that such questions provide an explicit structure to the individual in order to help them explore the issue using the questions as cues which in turn may encourage diverse responses and facilitate understanding (see e.g., Annesley, 2010; Paterson, 2006). The six hats tool is similar in that it refers to six distinctly coloured hats that emphasise a particular style or approach to thinking. For instance the yellow hat encourages the individual to focus on the positive issues whilst the black hat forces the individual to think about the negative consequences or risks (see, de Bono, 2009). The underlying rationale for this tool is that it provides an explicit framework to scaffold or facilitate creative thinking (see, Rizvi, Bilal, Ghaffar, & Asdaque, 2011).

It should be emphasised that there is nothing particularly special about these tools and the role they play in problem construction performance. They were selected for a number of reasons. First, is the simple pragmatic stance of having to begin the assessment of such tools somewhere and that the Six Thinking Hats is a well-known and popular tool that has been in circulation for some time (see, de Bono 2009). Given the six elements of this tool the Six Good Men, which also contains six elements, provides a good control/alternative. Nevertheless, it should be made clear that whilst we are focusing here on the Six Good Men and the Six Thinking Hats this does not preclude many of the other tools from potentially showing beneficial effects on problem construction performance (see e.g., Kurtzberg & Reale, 1999). Furthermore, and potentially more importantly, we wanted to know whether the problem construction benefit previously shown for these tools was simply the result of a placebo effect. For instance, we found that when used on a problem construction task, both tools proved to be more effective compared to a no-intervention control group (Vernon & Hocking, 2014). However, whilst suggestive differences were evident in effect sizes between the two interventions there were no clear differences between them. Given the fact that the control group were not given a tool to use it could be that use of a tool benefits a user through repetition, because the tool encourages six iterations, or placebo, because the tool promotes improvement through the strength of belief. The idea of a placebo influencing behaviour is widely documented in the scientific literature and a variety of evidence is available showing that an individual’s expectation can have a dramatic effect on behaviour (see e.g., Moseley et al., 2002). Hence, it may be that when given a tool to use on a problem construction task participants naturally expect their performance to improve. Furthermore, participants’ level of
motivation was not measured and those given a tool to use may have, as a consequence of using the tool, become more motivated to complete the task, which could also account for the benefit shown by those using a tool as motivation has been shown to be a key factor in creative performance (Amabile, 1983, 1996; Fasko, 2001; Sternberg & Lubart, 1999). Hence, to ascertain more precisely whether these two tools are capable of eliciting a beneficial effect on problem construction performance we compared performance on the two experimental interventions (i.e., *six men, six hats*) to a placebo intervention whilst simultaneously measuring participant motivation. Additionally, this placebo tool comprised of six elements to control for any potential iteration bias.

The placebo intervention developed for this study was called *brain breathing* and is based on the plausible links established between breathing influencing brain activity (e.g., Takahashi et al., 2005), in particular the alpha electroencephalographic frequency range which has been shown to be associated with creativity (see, Fink & Neubauer, 2006). The brain breathing technique simply requires participants to close their eyes, take three in-breaths and three out-breaths, and then open their eyes and note down any ideas that have occurred to them. Having a placebo tool that is comprised of six elements (i.e., 3 in-breaths and 3 out-breaths) helps to control for any potential iteration bias. In addition, it was thought that reference to a technique that directly involved the ‘brain’ would tap into the seductive allure of brain based explanations (see, Weisberg, Keil, Goodstein, Rawwon, & Gray, 2008).

Thus, the aim of this study was to compare the effectiveness of each experimental tool to that of a placebo-intervention on the same problem construction task. On the basis of evidence showing the facilitative effect of effortful, structured thinking (e.g., Reiter-Palmon & Robinson, 2009), and our previous research suggesting that the *six men* and the *six hats* are useful we predict that participants using either experimental technique would exhibit improved problem construction ability compared to the placebo-intervention control group. However, it is not clear at this stage whether any differences in problem construction ability would emerge between the two experimental techniques.
Method

Participants

One hundred and eighteen participants (22 male; 96 female) aged 18y to 35y (mean age 19.5y) took part in the study during an undergraduate psychology induction session. Participants were randomly allocated to one of three groups with each group focusing on the use of a specific tool (40-six hats, 40-six men, 38-brain breathing). All participants were volunteers and it was made clear that they were free to withdraw at any time and have their data removed/destroyed.

Materials

The study was conducted using specifically constructed workbooks. Each workbook recorded demographic (name, age, gender) information and contained seven self-report questions created using a 5-point Likert response scale. The first two assessed the participant’s views on creativity (Q1: how creative do you think you are? Q2: How important to you think creativity is in life?). The remaining five were used at the end of the study to obtain feedback on participants motivation to complete the task (Q3: How motivated were you to complete the task?), familiarity with the allotted technique (Q4: Have you ever used the specified technique before?), and feedback on use of the technique (Q5: How easy/difficult did you find it to use the technique? Q6: How useful did you find it to use the technique? Q7: How likely is it that you would use the technique in the future?). The workbooks also contained an introduction to the technique (i.e., six hats; six men; brain-breathing) along with an example problem chosen specifically to be relevant to the students (How can I improve my academic grades?) with examples of how the technique could be used to help explore and understand the problem. This was followed by a brief explanation of problem finding and the focus on restating the problem to aid understanding and finally the problem used in the main part of the study, which was the same for all participants: ‘I am in a new city and need dinner’, which was taken from Paletz and Peng (2009). Beneath this was a grid containing 18 boxes for the participants to write in their restatements, with one box per restatement.

---

1 This problem construction study represented part of a larger study, the remainder of which will be reported elsewhere.
Design

The study used a between participants design with a single factor of Group with three levels (six hats; six men and brain-breathing). Four dependent measures were used to assess problem construction performance. The first was fluency, which referred to the number of problem restatements (see, Fontenot, 1993). The second, quality/usefulness, captured the degree to which the problem restatements were likely to result in a logical/workable approach to the situation, and was scored on a five point Likert scale from 1 (very low quality) to 5 (very high quality) (see, Mumford et al., 1996). The third measure was flexibility which referred to the number of conceptual categories into which the restatements could be classified (after, Sowden, Clements, Redlich, & Lewis, 2015). The fourth and final measure was originality and assessed using the formula: originality = 1- the frequency of a given restatement / total sample size (after, Sowden et al., 2015; Zenasni & Lubart, 2009).

Procedure

There were four timed phases in the experiment and each participant completed them in the same order. Phase 1, which took 5 minutes, was used to introduce the study as a ‘creative problem construction task’ and provided information on the nature of the study, as well as obtaining informed consent. Phase 2, which also took 5 minutes, required participants to enter their demographic information onto their workbook and read through the explanation of the technique and example given. In Phase 3 twenty minutes was allocated for participants to read through the brief explanation of problem finding and complete the main task by entering as many restatements to the posed problem as they could in the grid below. Following this, participants were given three minutes to complete Phase 4, which comprised the post-problem construction questions regarding motivation, familiarity with the technique, how easy/difficult it was to use the technique, how useful they thought using the technique was and whether they would consider using the technique in the future. Once completed the workbooks were collected, the participants were thanked and the two experimenters debriefed them regarding the aims and objectives of the study providing additional contact details should they wish to ask any further questions.

Results

Two independent raters blind to the aims/objectives of the study were used to code and rate all responses. Consistent agreement was obtained for responses to self-report questions and
the measure of fluency. For quality, where coded responses differed by more than one rating point in either direction (<21%), a third blind rater was brought in to arbitrate the decision. Inter-rater reliability as measured by intra-class correlations (Shrout & Fleiss, 1979) was 0.81. For flexibility, intra-class correlations between the number of conceptual categories identified by each rater was good at 0.70. This classification included categories such as, location of food; information gathering; use of technology; travel/transport issues and money/price. Whilst the intra-class correlations of the originality score were also good at 0.81 it should be noted that using the formulaic approach, as outlined above, to identify originality means that it is not simply that participants in one group produce a restatement that participants in another group do not, but in order for a particular group to obtain a higher originality rating it would mean that participants within this group produce the more unusual (i.e., more original) restatement more of the time. An example of this were restatements that focused on ‘location’, with those restatements that simply focused on locating a food provider, such as a restaurant, receiving a lower originality score (e.g., 0.6299) compared to restatements that focused on locating alternative sources of food (e.g., 0.8831). Descriptive statistics regarding responses to the initial questions on participants’ views of creativity are presented in Table 1. This shows that participants in each group rated their own creativity levels similarly, at around the mid-way point. However, they all rated the ‘importance of creativity in life’ as significantly more important (grand means of 2.96 and 4.02 respectively; t(118)=10.81, p<0.001, d=1.87).

Table 1 about here

To test the predictions that the structured interventions (i.e., six men; six hats) would lead to improved creativity performance relative to the placebo intervention (i.e., brain-breathing) a one-way analysis of variance (ANOVA) was carried out on each of the four creativity measures (i.e., fluency, quality, flexibility, originality) with orthogonal planned contrasts comparing performance of each intervention to placebo. For Fluency this led to a main effect of Group F(2,115)=9.385, p=0.001, Mse=11.58, $\eta^2=0.14$ with contrasts showing that those
using the six men technique produced significantly more restatements than brain-breathing (mean fluency 8.67 vs. 5.6 respectively; \( p=0.001, d=0.8 \)), whilst those using the six hats showed no difference relative to brain-breathing (mean fluency 6.05 vs. 5.60). Post-hoc analysis showed greater fluency for those using the six men compared to the six hats techniques (mean fluency 8.67 vs. 6.05 respectively; \( t(78)=3.41, p=0.001, d=0.8 \)). Analysis of Quality showed no main effect of Group \( F(2,115)=0.887, p=0.415, \text{Mse}=0.56, \eta^2=0.015 \) with contrasts showing no difference between those using the six men technique and brain-breathing (mean quality 3.48 vs. 3.27 respectively), and no difference between those using the six hats and brain-breathing (mean fluency 3.30 vs. 3.27 respectively). For Flexibility there was a main effect of Group \( F(2,115)=15.347, p=0.001, \text{Mse}=3.06, \eta^2=0.21 \) with contrasts showing that those using the six men technique produced restatements from more conceptual categories than brain-breathing (mean fluency 6.15 vs. 4.05 respectively; \( p=0.001, d=1.09 \)), whilst those using the six hats showed no difference relative to brain-breathing (mean fluency 4.65 vs. 4.05). Post-hoc analysis showed greater flexibility for those using the six men compared to the six hats techniques (mean fluency 6.17 vs. 4.65 respectively; \( t(78)=4.01, p=0.001, d=0.92 \)). Analysis of Originality revealed a main effect of Group \( F(2,115)=12.119, p=0.001, \text{Mse}=.001, \eta^2=0.174 \) with contrasts showing that those using the six men technique produced more original restatements compared to those using brain-breathing (mean originality .1097 vs. .0698 respectively; \( p=0.001, d=1.05 \)), and those using the six hats also produced more original restatements compared to brain-breathing (mean originality .0862 vs. .0698 respectively; \( p=0.048, d=0.51 \)). Post-hoc analysis also showed greater originality for those using the six men compared to the six hats techniques (mean originality .1097 vs. .0861 respectively; \( t(78)=2.896, p=0.005, d=0.67 \)). Thus, structured approaches improved fluency, flexibility and originality scores but not quality.

Descriptives statistics of responses to the post-restatements questions regarding the difficulty of the task, the use of the technique and possible future use are shown in Table 2.

______________

Table 2 about here

______________
A one-way ANOVA conducted on participants ratings of their motivation levels showed no main effect of Group F(2,115)=0.241, p=0.786. In terms of how easy/difficult it was to use the technique there was a main effect of Group F(2,113)=7.047, p=0.001. Further comparisons using a Bonferroni correction showed that those using the brain-breathing technique found this easier than those using the six hats (mean ease of use 3.57 and 2.80 respectively; p=0.001, $d=0.78$) and that those using the six men found it easier than those using the six hats (mean ease of use 3.34 and 2.80 respectively; p=0.037, $d=0.62$). Analysis of how useful participants found using the technique produced a main effect of Group F(2,113)=9.61, p=0.001. Further comparisons using a Bonferroni correction showed that those using the six men technique found this more useful than those using the six hats (mean usefulness 3.73 and 3.17 respectively; p=0.03, $d=0.62$) and those using brain breathing (mean usefulness 3.73 and 2.78 respectively; p=0.001, $d=1.07$). Finally, analysis of how likely it is that participants would use the technique in the future produced a main effect of Group F(2,113)=8.678, p=0.001. Further comparisons using a Bonferroni correction showed that those using the six men technique would be more likely to use this technique again in the future compared to those using the six hats (mean usefulness 3.78 and 3.10 respectively; p=0.01, $d=0.71$) and those using brain breathing (mean usefulness 3.78 and 2.84 respectively; p=0.001, $d=1.01$).

**Discussion**

Use of the six men tool led to greater fluency, flexibility and originality compared to both the six hats and the placebo intervention. Whereas use of the six hats tool led to more original responses compared to the placebo intervention. There was no difference in reported motivation between the three conditions. However, those using the six men and the brain-breathing tools found these easier to use compared to those using the six hats. In addition, those using the six men found this tool to be more useful and reported that they would be more likely to use it again in the future.

These findings support our previous work showing that use of a tool that provides explicit structure can aid problem construction ability (Vernon & Hocking, 2014) and are consistent with others who have shown that training can improve creative performance (Feldhusen & Clinkenbeard, 1986; Mumford et al., 1994; Reiter-Palmon, Mumford, Boes, & Runco, 1997). It also extends our previous work to show that such effects can be elicited on a
younger age group. Furthermore, the findings indicate that the benefit elicited from using a tool is not simply the result of a placebo effect or the result of differential motivation levels as a result of using a tool, but that use of a specific tool helps to provide explicit structure that may open up new ways of thinking for the individual and help them to shift between perspectives which in turn helps them to engage in the problem construction process (see, Liu & Schonwetter, 2004).

It is interesting to note that overall the six men tool elicited a more robust effect on measures of creative performance. In part this may be because the tool may be easier to use and/or apply, which would be consistent with the findings reported here. The six men simply refers to six questions that participants will be familiar with and they are easily read and recalled. In contrast, the six hats represents a more conceptually rich tool with each hat involving a particular type of thinking (see, de Bono, 2009). Indeed, participants in this study found the six hats tool to be the most difficult to use. Given that the study relied on participants reading a brief explanation of the relevant tool along with an example to provide some insight into its use, it should come as no surprise that the tool that was easy to learn proved to be the more effective one. It may well be the case that with additional practice and/or direct instruction on the use of the six hats tool, its effectiveness would improve. While speculative, such a possibility would be consistent with the findings of others who have found that additional practice is often required to elicit a clear effect when training creative problem solving performance (see e.g., Daniels, Heath, & Enns, 1985; Wang & Horng, 2002) and that directed study can have a greater impact than self-directed study (Hunsaker, 2005).

The fact that use of the six hats produced a less robust effect could also be the result of the order in which the hats were used. Of course, it may be that for some problems some hats are more useful than others and not all hats may be required in all situations. Or that the particular order is irrelevant so long as all six are used to provide an overview of the problem. Or that individuals should spend more time metaphorically wearing some hats compared to others. Unfortunately there is no clear consensus in the literature on this issue. For instance, Pohl (1994) suggests that processes associated with exploring and inventing could use a sequence which begins with the blue hat, followed by the green and red hats. In contrast Paterson (2006) suggests beginning with the yellow hat to ‘set the stage’ (p.11) followed by green and red hats. Whereas de Bono (2009) outlines the six thinking hats in the order of white, red, black, yellow, green, and blue. Given the lack of data and the high number of
possible permutations of the six hats (720) it may not be possible to examine them all. Nevertheless, future research could manipulate the order of the hats to ascertain what, if any, effect this may have on creative performance.

An additional point is that the effectiveness of the tool may be linked to the complexity of the problem. Here the given problem (i.e., I’m in a new city and need dinner) was reasonably clear and the participants task was to construct as many additional problem restatements within the given timeframe (cf., Reiter-Palmon et al., 1997). With this in mind it is no surprise that the easy-to-use tool proved most beneficial, though it should be noted that use of the six hats also led to greater originality relative to the placebo group. Nevertheless, a more ambiguous and/or ill-defined problem would be more difficult, requiring greater cognitive effort and as a consequence performance on a task utilising such a problem may be more influenced by a tool that provides greater conceptual information (see, Paletz & Peng, 2009). Hence it is possible that comparing the two tools (i.e., six men and six hats) on a problem construction task utilising a more complex problem may elicit a different pattern of effects. In this instance use of the six hats tool, which provides more conceptual information could help to scaffold thinking in more direct ways, which in turn may have a greater effect on performance. Though speculative, such a proposal would be consistent with the view that techniques that provide more structure can have a greater effect on performance (Paletz & Peng, 2009; Scott, Leritz, & Mumford, 2004a). Alternatively, or in addition to this complexity effect, it might be that there is a better match between the content of the tool and the problem scenario in question. Whether this is because of common characteristics (such as the six men and ‘new city’ problem arguably sharing a theme of orientation) or because the tool happens to be a route to improved performance (cf. the Amusement Park Model of Baer & Kaufman, 2005), are not clear from our data and as such remain the domain of future research.

An issue not explored within this study but relevant to the field of training within CPS is the duration of any effects elicited by such training. Others have suggested that the benefits seen in training may persist over long periods of time (Feldhusen & Clinkenbeard, 1986). Hence, future research could incorporate follow-up assessments to ascertain the possible long-term benefits of such training.

Finally, similar to our previous study we found no differences in the ‘quality’ of responses across the three groups. Again this should not be taken as evidence that the use of
such tools is insufficient to elicit any beneficial effects on the quality of the restatements. Because our study focused on the putative generation stage of CPS, it provided no opportunity for participants to engage in an evaluative stage where poorer or less useful ideas could be dropped. Such elimination would likely increase the overall quality rating for all techniques and, given a consistent elimination threshold (i.e. an unbiased judgement about the elimination of restatements), we should see differential effects of quality by group if such quality differences are true. Furthermore, feedback from our coders indicated that despite our attempts to fine tune this variable and adopt a more concise definition (after, Mumford et al., 1996) they found this very difficult to code. It may be that more time is needed to train our coders or that more precise guidelines need to be provided indicating what a ‘good’ quality restatement would look like.

In conclusion, we show here that the benefits from using a structured thinking tool to enhance problem construction performance are robust and not the result of a placebo effect or due to differential motivation levels. The tool that was easier to learn proved, in this instance, to be the more effective. However, further work needs to be done to clarify the role of directed instruction and the relationship of the tool to the problem.
References


Table 1. Showing mean responses, with standard deviations (SD), to initial self-report questions on a scale from 1 (not at all) to 5 (very) from participants within each group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>How creative do you think you are?</th>
<th>How important do you think creativity is in life?</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Hats</td>
<td>3.10 (1.25)</td>
<td>3.92 (0.79)</td>
</tr>
<tr>
<td>6 Men</td>
<td>2.90 (0.84)</td>
<td>4.17 (0.91)</td>
</tr>
<tr>
<td>Brain-Breath</td>
<td>2.89 (1.01)</td>
<td>3.97 (0.91)</td>
</tr>
</tbody>
</table>
Table 2. Showing mean responses, with standard deviations (SD), to post-restatement task self-report questions on motivation (1=not at all; 5=extremely motivated), task difficulty (1=extremely difficult; 5=extremely easy), usefulness (1=not at all useful; 5=extremely useful) and likelihood of using the technique in the future (1=not at all likely; 5=extremely likely).

<table>
<thead>
<tr>
<th>Condition</th>
<th>How motivated were you?</th>
<th>How easy/difficult to use the technique?</th>
<th>How useful did you find the technique?</th>
<th>How likely to use the technique in the future?</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Hats</td>
<td>3.10 (0.84)</td>
<td>2.80 (0.91)</td>
<td>3.17 (1.03)</td>
<td>3.10 (1.15)</td>
</tr>
<tr>
<td>6 Men</td>
<td>3.22 (1.18)</td>
<td>3.34 (0.85)</td>
<td>3.73 (0.76)</td>
<td>3.78 (0.77)</td>
</tr>
<tr>
<td>Brain-breathing</td>
<td>3.05 (1.33)</td>
<td>3.57 (1.05)</td>
<td>2.78 (1.01)</td>
<td>2.84 (1.10)</td>
</tr>
</tbody>
</table>