A Cognitive Procedure for Representation Change in Verbal Insight Problems

John Patrick, Afia Ahmed, Victoria Smy, Helen Seeby, and Katie Sambrooks
Cardiff University

The aim of this study was to develop a novel cognitive procedure for operationalizing how the re-encoding and constraint relaxation, suggested by representational change theory (RCT) (Ohlsson, 1992, 2011), can effect representational change in verbal insight problem solving, thus circumventing the constraints imposed by past experience. Some participants were trained in using an evaluative cognitive procedure that aimed to facilitate the identification of any inconsistency between the participant’s interpretation of the problem and the problem statement, and thus cue the re-encoding proposed by RCT. In Experiment 1, participants were randomly allocated to training, practice, or a no-training control condition, and were subsequently tested on 7 verbal insight problems. Concurrent verbal protocols were collected and analyzed to identify problem solvers’ proposed hypotheses and also to assess whether problem solving behavior changed in line with the training. Inconsistency identification training, rather than practice or no training, improved solution rate across novel problems and resulted in more paraphrasing and questioning of the problem statement, and a modest increase in participants’ reflection on their problem solving. Results from Experiment 2 indicated that this improvement in representation change through training was not due to increased awareness of the nature of verbal insight problems but rather training in identifying inconsistencies between the problem statement and a person’s interpretation of it. Experiment 3 revealed that the performance improvement with training was sustained after a delay of 48 hr. Theoretical and methodological issues are discussed.

Keywords: representation change, insight problem solving, training, transfer

The aim of this study was to elaborate and use representational change theory (RCT) (e.g., Kaplan & Simon, 1990; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Ohlsson, 1992, 2011), particularly the mechanisms proposed by Ohlsson, to develop a cognitive procedure that could be trained and used to improve the solution of verbal insight problems. This would then obviate the need for the problem solver to have to experience repeated negative feedback, suggested by RCT, in order to reduce the activation of an inappropriate conceptualization of the problem and effect representation change. Breaking an impasse in an insight problem is important yet difficult because it is the result of a well-learned implicit constraint or assumption that is usually correct. Ohlsson (1992) proposed that representational change can be effected through the cognitive mechanisms of re-encoding and constraint relaxation, although the potential role of training was not elaborated. A recent study by Patrick and Ahmed (2014) focused on the role of training in RCT, demonstrating how training with heuristics can improve the solution rate of a category of verbal insight problems, rather than practice or no training, improved solution rate across novel problems and resulted in more paraphrasing and questioning of the problem statement, and a modest increase in participants’ reflection on their problem solving. Results from Experiment 2 indicated that this improvement in representation change through training was not due to increased awareness of the nature of verbal insight problems but rather training in identifying inconsistencies between the problem statement and a person’s interpretation of it. Experiment 3 revealed that the performance improvement with training was sustained after a delay of 48 hr. Theoretical and methodological issues are discussed.

Keywords: representation change, insight problem solving, training, transfer

Training Studies of Representation Change

A major issue in studies of training for verbal and visuospatial insight problems is the degree of specificity of the training. Many studies have provided specific training oriented to the solution of one particular problem, such as the four-dot problem (e.g., Weisberg & Alba, 1981, Experiment 4) and the nine-dot problem (e.g., Chronicle, Ormerod, & MacGregor, 2001, Experiment 3; Kershaw & Ohlsson, 2004; Weisberg & Alba, 1981, Experiment 2). A general criticism of such specific training is that the novelty of the problem may be diminished with the solution path being prompted. A similar problem arises in studies in which specific hints are used (e.g., Bumscham & Davis, 1969; Weisberg & Alba, 1981, Experiment 1). Indeed it might be argued that, by giving away some details of the problem solution in the training or through a hint, there may no longer be a need for representation change, in
contrast to participants without such extra information (G. Jones, personal communication, March 7, 2013). Consequently, it is important that any training intervention underpins positive transfer to a range of different verbal problems, and representation change can be verified and attributed to the training.

Relatively few studies have investigated how a cognitive procedure can be designed and trained to facilitate the solution of not only one problem for which the training was designed but also other previously unseen problems. The studies by Ansburg and Dominowski (2000); Chrysikou (2006); Dow and Mayer (2004); Walinga, Cunningham, and MacGregor (2011); and Wicker, Weinstein, Yelich, and Brooks (1978) are of particular interest in this respect because training did not directly cue the nature or category of the test problem(s), and the training was intended to facilitate representation change for a variety of verbal insight problems with different constraints. The study by Chrysikou (2006), using a form of creativity training that involved attempting to break the stereotypical categorization of critical objects, revealed that problem solving improved to approximately 60% irrespective of whether the training included the problem-specific object or not. This result is interesting and surprising as the training effect did not appear to be limited to functional fixedness problems (E. G. Chrysikou, personal communication) for which it was particularly suited. The studies by Wicker et al. (1978) and Ansburg and Dominowski (2000) adopted similar approaches by providing some strategic instruction coupled with practice and solution feedback. In Experiment 2 of the study by Wicker et al. (1978), participants who received “reformulation” training were exhorted not to define the problem too narrowly or make unnecessary assumptions about it and were given practice at doing this with eight problems. Solution rate was best for this reformulation strategy (63%) in comparison to two control conditions. However, this finding is difficult to assess because the reformulation strategy was not specified in detail, and a list of the training and test problems was not given, making it problematic to judge the transfer relationships between them. Also a 5-point scoring scheme was used to assess each test solution, although the criteria for this scheme were not made explicit. The study by Cunningham and MacGregor (2008) also involved strategic instruction and practice and revealed no effect of training with matchstick arithmetic problems on the solution of verbal insight problems.

Ansburg and Dominowski (2000), in a series of five experiments, investigated the effects of various forms of training on the solution of novel verbal insight problems. In Experiments 1 and 3, solution rates for training were 60% and 57% compared with control conditions of 34% and 35%, respectively. However, the training contained many diverse elements, including (a) reading strategic instructions that emphasized reading a problem carefully, interpreting it in alternative ways, and not getting stuck on the most obvious interpretation; (b) practicing and, when unsuccessful, rereading the problem twice, paraphrasing the problem; (c) explaining, if still unsuccessful, why the provided solution is correct; and (d) comparing and identifying the procedural similarities among practice problems. It is difficult to disentangle the effect of these different training manipulations. In common with the studies by Wicker et al. (1978) and Cunningham and MacGregor (2008), the strategic instructions concerned trying to think in a different way. However, these studies did not provide any specific cognitive procedure in training of how to accomplish this, and the training solution was not grounded in any theoretical framework concerning the nature of an impasse and how it may be overcome.

**RCT and Designing a Cognitive Procedure**

This paper focuses on verbal insight problems that, arguably, have received less research attention than visuospatial problems. Thus, for example, if we are asked: “A man pushed a car. He stopped when he reached a hotel at which point he knew he was bankrupt. Why?” we immediately hit an impasse because we automatically and inappropriately represent the situation as involving a real car rather than a car on a monopoly board. Similarly, 1,988 pennies are worth more than 1,983 pennies, not because of their apparent dates that act as constraints, but because there are more of them!

In the original account of RCT, Ohlsson (1992) explained how normal cognitive processing in problem solving can lead to an impasse and also how it can be overcome by the mechanisms of re-encoding and constraint relaxation that effect representation change. In Ohlsson’s (2011) more recent contribution, labeled redistribution theory, he elaborated a spread and level of activation account of retrieval from memory that was determined by past experience. So when confronted with an incomplete problem description, such as *Push That Car* problem, an automatic interpretive and constructive process takes place with an inappropriate representation activated on the basis of past experience such that, in the absence of further information, the car is assumed to be a real car. Ohlsson (2011) posited that this inappropriate representation may change as the problem solver receives increasing amounts of negative feedback during the impasse phase, resulting in some activation being subtracted from the parts of the cognitive structure found to be unsuccessful in generating a solution. This activation is then redistributed across other parts, such that gradually there will be sufficient activation elsewhere to provide a revised solution space. However, there is no guarantee that this process will identify the appropriate representation change with a revised solution space or that this will be accomplished in any subsequent revisions. An alternative and more direct approach is to try to implant, through training, a cognitive procedure that will have sufficient activation strength to be recalled during an impasse and may provoke the necessary re-encoding and representation change. This was the goal of the present article.

There are various issues to be considered in the design of such a cognitive procedure. Our first assumption is that it will not be possible to develop a procedure to inhibit the problem solver from generating an inappropriate representation after initial reading of the problem. The reason is that researchers generally consider that this initial incorrect representation is the result of an automatic and largely unconscious process in insight problem solving (e.g., Jones, 2003; Knoblich et al., 1999; Ohlsson, 1992, 2011; Ollinger, Jones, & Knoblich, 2008; Ormerod, MacGregor, & Chronicle, 2002). This perspective is consistent with evidence from the areas of text comprehension and reading that emphasize the automatic generation of some aspects of meaning that are strongly guided by past experience (e.g., Graesser, Millis, & Zwaan, 1997; Kamide, Altmann, & Haywood, 2003; Seidenberg, 2005; Trueswell, Tanenhaus, & Kello, 1993).
Consequently, any cognitive procedure will have to operate after an inappropriate representation is generated. Also, as previously stated, this cognitive procedure has to be capable of solving a range of verbal insight problems, and it therefore cannot involve any specific heuristics tailored to overcome the constraints of one problem or even a category of problem, as used in the studies, for example, by Adeyemo (2003) and Patrick and Ahmed (2014). Instead the cognitive procedure has to capitalize on some commonality in terms of what happens when solving verbal insight problems. A common feature of verbal insight problem solving is that a person’s default and inappropriate interpretation will inevitably contain some detail or unwarranted assumption that is inconsistent with the problem specification. For example, in the *Coins of the Realm* problem, it is assumed that 1988 refers to a date, although this is not actually specified in the problem statement. If such a discrepancy could be identified by the problem solver, then it could become the basis for changing an inappropriate representation into an appropriate one for solving the problem. Therefore, in the present study, the problem solver is trained to use a cognitive procedure that focuses on trying to identify any inconsistency between a person’s interpretation of the problem and the actual problem specification. Identifying such an inconsistency could provide not only the trigger for re-encoding, but the nature of the inconsistency could indicate how re-encoding should take place to solve the problem. This should improve solution rate but would not guarantee solution because either the “correct” inconsistency may not be spotted or the search with the revised representation may still be unsuccessful.

A similar approach to improve problem solving was used in a study by Patrick et al. (1999) in which nuclear power plant operators were trained to diagnose unusual multiple faults that otherwise they found near impossible to solve. These operators were very experienced and found it difficult to shift from their normal default reasoning that involved a single rather than a multiple fault. Consequently, training was developed that provided operators with awareness of what might block problem solution and also practice at trying to identify the inconsistency between the actual symptoms and those that were associated with their incorrect hypothesized fault. Operators were encouraged to make this evaluation if they were unable to diagnose the problem. Any inconsistency that was identified could then serve as a cue to extend a stereotypical single-fault hypothesis into a more unusual multiple one by generating a further fault that could be responsible for the remaining symptoms that were inconsistent and unaccounted for by the initial hypothesis. Using such a mechanism was feasible because consistency checking has been found to be a ubiquitous process in hypothesis generation that takes place spontaneously even before people are explicitly asked to consider the plausibility of their hypothesis (Fisher, Gettys, Manning, Melehe, & Baca, 1983).

**Experiment 1**

In the present study, training focused on developing an evaluative cognitive procedure aimed at identifying any inconsistency, including any unwarranted assumption or detail, between the person’s interpretation or understanding of the problem situation and the formal problem specification. Any identified inconsistency may then cue re-encoding and effect an appropriate change in representation. In order to achieve this, trainees were first made aware of how an habitual interpretation can block solution of an insight problem and the relevance of the training for overcoming such difficulty (Brown, Campione, & Day, 1981; Campione & Armbuster, 1985; Gick & Holyoak, 1987). In addition, from the insight training studies by Wicker et al. (1978) and Ansburg and Dominowski (2000), and the general training literature, a powerful training design principle is to give practice with feedback at doing in training exactly the same cognitive processing as would be required in the testing or transfer situation (e.g., Anderson, 1987; Needham & Begg, 1991; Patrick, 1992). These principles guided the design of training.

A further training requirement is that the cognitive procedure is comprehensively applied to all parts of the problem specification in order that the possibility of generating the solution is maximized. Therefore, an additional feature of the cognitive procedure in the present study is that it encouraged an iterative process of searching. Participants were trained to select one part of the problem and attempt to spot any inconsistency between the problem statement and their interpretation, and if none was found, then a further part of the problem was selected and scrutinized.

To evaluate the effectiveness of the cognitive procedure, participants in the training condition were compared against those in two control conditions: a practice condition, in which participants received solution feedback during training, and a no-training control condition. Concurrent verbal protocols were collected, transcribed, and analyzed in order to identify correct solutions and collect supplementary qualitative evidence of the effect of the trained cognitive procedure on problem solving as advocated by Ash, Cushen, and Wiley (2009). It was hypothesized that training would result in an increase in solution rate; a reduction in the generation of hypotheses that were inconsistent with the problem statement; more rereading or paraphrasing of all or part of the problem statement; more questioning of the problem statement; and more self-questioning or reflection on problem solving strategy.

**Method**

**Participants.** Thirty-six psychology students from Cardiff University took part in this study. Ages ranged between 18 to 35 years ($M = 19.50$ years, $SD = 2.80$). Participants were randomly allocated to one of three conditions: inconsistency identification training ($n = 12$), practice ($n = 12$), and no training ($n = 12$).

**Training program and problems.** Inconsistency identification training comprised two parts (see Appendix A). The aim of the first part was to provide general awareness of the nature of verbal insight problems and how an incorrect interpretation can block solution. Two examples were given of how this blocking effect might operate in the *Push That Car* and the *Coins of the Realm* problems. In the second part of training, the iterative cognitive procedure was described, and participants practiced using it to identify any inconsistency between the problem specification and their interpretation of any part of the problem. This training used the *Sid Shady*, *Barney Dribble*, and *Coffee* problems with solution feedback from the experimenter (for more detail, see Appendix A).

The training and test problems together with their constraints are described in Appendix B. Each of seven test problems was presented on a separate sheet of A4 paper. They were the *Pear Tree*, *Steady As She Goes*, *Coins of the Realm*, *Push That Car*, *Sid Shady*, *Barney Dribble*, and *Coffee* problems.
Participants were individually tested and/or trained. Participants in the inconsistency identification training condition were informed of the general nature of verbal insight problems (i.e., raising awareness in Part 1 of the training, see Appendix A). They were then instructed in the iterative cognitive procedure and practiced using it (see Part 2 of Appendix A). Participants in the practice condition received no training. Participants in the no-training condition completed the word association task (Cohen, 1975) as a filler task before being tested. Participants in all conditions then attempted the seven test problems that were presented in random order for each participant, with a time limit of 4 min for each, as pilot testing revealed that this was sufficient time to solve the problem if they were going to do so. After completing each test problem, participants were required to rate their familiarity with that problem on a 5-point scale, although this did not result in the rejection of any problems or participants in Experiment 1. Participants were not given solution feedback and were asked not to reveal information about the experiment to others. The durations of the inconsistency identification training, practice, and the word association control task conditions were approximately 55, 40, and 40 min, respectively.

During the test phase, concurrent verbal protocol data were collected. Participants were instructed in how to “think aloud” during problem solving, and to facilitate this, practice exercises were given involving a multiplication problem, calculating the number of windows in the participant’s house, and naming 20 animals, as recommended by Ericsson and Simon (1980, 1993). If participants fell silent during problem solving, the experimenter used two nondirective prompts (“What are you thinking?” and “Please keep talking”). Verbalizations were recorded continuously.

Analysis of verbal protocols and intercoder reliability. All coders were blind with respect to what condition was being coded. Three coders (the second author and two researchers) individually identified from each transcribed protocol the hypotheses proposed for each test problem. Hypotheses were identified because they are important features of cognitive processing and can indicate whether a participant has been able to break the problem constraint. Practice at coding was undertaken on one verbal protocol that was subsequently excluded. A total of 593 hypotheses were initially identified. Of these hypotheses, 574, or 96.8% were identified by all three coders, four, or 0.8% were identified by two coders, and 15, or 2.5% were identified by one coder. After discussion of these discrepancies, 18 hypotheses were rejected, making the final total 575.

Subsequently, the same coders independently categorized these hypotheses as correct, inconsistent (with the problem statement), or consistent but incorrect. Inconsistent hypotheses were so named because they contradicted some information given in the problem statement. Consistent but incorrect hypotheses were those that did not contradict any information in the problem specification but nonetheless were incorrect. Of the 575 hypotheses, it was found that 574, or 99.8% were categorized by all three coders using the same category, one, or 0.2% by two coders, and none by only one coder. The Perreault and Leigh (1989) reliability indices for the three pairs of coders were at the maximum of 1.00.

Further analysis of protocols aimed at identifying changes in thinking attributable to the inconsistency identification training procedure. Transcribed protocols were analyzed by three researchers, different to those used above, to identify instances when participants (a) reread or paraphrased all or part of the problem statement, (b) questioned some aspect of the problem statement, and (c) self-questioned or reflected on their problem solving. The frequency of behaviors falling into the first two categories was predicted to increase as a direct consequence of the trained cognitive procedure. The last category was included because it was considered that an indirect consequence of training might be for participants to develop a more general reflective and questioning approach to problem solving strategy. A total of 814 segments were initially identified in these three categories, of which 728 (or 89.43%) were identified by all three coders, 28 (or 3.44%) were identified by two coders, and 58 (or 7.13%) by one coder. After discussion of these discrepancies, 58 segments were rejected, making the final total 756. Next, the coders categorized the 756 segments into one of the three training-related behaviors. The Perreault and Leigh (1989) reliability indices between the three pairs of raters were satisfactorily high at 0.98, 0.97, and 0.99.

Results and Discussion

The first analysis aimed to establish whether participants initially generated an inappropriate representation after reading the problem. This was because a prerequisite for the success of the inconsistency identification cognitive procedure was that such a misrepresentation was generated initially; otherwise, there was no opportunity for an inconsistency to be identified. A second reason was that in order to claim that representation change has taken place, we needed to establish that initial thinking was characterized by misrepresentation and that subsequent thinking of trained participants was not. We therefore analyzed data concerning the nature of first hypotheses proposed and also coded initial verbal comments before the first hypothesis was proposed in terms of whether they reflected a misrepresentation. These data will understate the extent of an initial misrepresentation because it is always possible that a fleeting misrepresentation is initially generated that changes quickly to an appropriate representation for the problem, leaving no verbal evidence of this change. We analyzed data from all three conditions of Experiment 1 as there was no reason for there to be any differences in the frequency of initial misrepresentations of the problem among the three conditions. In Experiment 1, 12 participants in each of three conditions tackled seven problems, creating 252 opportunities for examining evidence that their initial problem solving was dictated by a misrepresentation of the problem. (All hypotheses in the protocols of Experiment 1 were reliably identified and coded concerning whether they were generated under a misrepresentation, as reported in the previous section.) We therefore collated data concerning whether the first hypotheses proposed were consistent with a misrepresentation or not. In 183 (73%) of problems, the first hypothesis involved a misrepresentation of the problem. The verbal protocols associated with the remaining 69 problems were coded independently by the third and fourth authors concerning whether there was any verbal evidence in the initial part of the transcript, before a hypothesis
was generated of a misrepresentation of the problem. In 30 out of 69 situations (with Perreault Leigh reliability index of .90), evidence of a misrepresentation was found. Thus, in total including the hypothesis data, 213 out of 252 (85%) of problem solving situations, there was evidence from the verbal protocols of an initial misrepresentation of the problem by participants in all three conditions. There was no difference among the three conditions in either the number of first hypotheses involving a misrepresentation, $F(2, 33) = 1.56, MSE = 2.29, p = .22, \eta^2 = .09$, or the number of verbal references indicative of an initial misrepresentation, $F(2, 33) = 2.29, MSE = .99, p = .12, \eta^2 = .12$. These analyses provide strong evidence that, after reading the verbal insight problems, there is a pervasive tendency to generate an initial misrepresentation of the problem. This is the case for all participants irrespective of condition.

The next step involved examining whether the trained inconsistency identification procedure resulted in representation change with participants in this condition, breaking the problem constraint and improving solution rate. Participants who received training in inconsistency identification produced the highest solution rate of 68% (57/84 problems correct) compared with 39% (33/84 problems correct) and 30% (25/84 problems correct) solution rates of those in the practice and no-training conditions, respectively (see Figure 1). An analysis of variance (ANOVA) conducted on the frequency of correct hypotheses revealed a significant difference among conditions, $F(2, 33) = 12.85, MSE = 1.80, p < .001, \eta^2 = .44$. Bonferroni comparisons revealed that participants trained in inconsistency identification solved more problems than those in both no-training and practice conditions ($p < .001$ and $p < .01$, respectively), and there was no difference in the solution rate of participants in these two latter conditions ($p > .05$). On average, the inconsistency identification training improved solution rate by approximately 33% compared with the practice and control conditions.

Further analyses investigated qualitative and quantitative effects of inconsistency identification training on aspects of problem solving. First, participants in this training condition were expected to propose fewer incorrect hypotheses that were contradicted by information in the problem statement, although this would constitute an indirect rather than a direct effect of training. An ANOVA revealed that there were no differences among participants in the three conditions in the generation of inconsistent hypotheses, $F(2, 33) = 1.45, MSE = 2.47, p = .24, \eta^2 = .08$. Next, two problem solving behaviors, directly related to training, were examined in order to determine whether trained participants exhibited more of them, and therefore their improved solution rate was likely to be a consequence of the cognitive procedure implanted during training. These training-related behaviors were rereading or paraphrasing some part or the whole of the problem statement, and questioning some aspect of the problem statement. Trained participants engaged in considerably more of these behaviors than participants in the no-training and practice conditions (see Table 1, and also see an example coded protocol in Appendix C). An ANOVA revealed a significant difference among conditions in the amount of rereading or paraphrasing all or parts of the problem statement, $F(2, 33) = 5.78, MSE = 69.87, p < .01, \eta^2 = .26$. Bonferroni comparisons confirmed that trained participants engaged in significantly more of these behavior than participants who were not so trained ($ps < .01$). An example of restatement for the Pear Tree problem is, “So the farmer California owns a beautiful pear tree. He supplies the fruit to a nearby grocery store. The store owner called the farmer to see how much fruit is available for him to purchase . . . .”

Similarly, trained participants questioned some aspect of the problem statement over twice as often as untrained participants (see Table 1). An ANOVA revealed a significant difference among conditions in the frequency of this training-related behavior, $F(2, 33) = 11.17, MSE = 7.17, p < .001, \eta^2 = .40$, with Bonferroni comparisons confirming that trained participants engaged in significantly more of this verbal behavior than participants in the practice and no-training conditions ($ps < .001$ and .01, respectively). An example of questioning for the Professor Bumble problem is, “It doesn’t say here what he was travelling in.”

Finally, we considered that an indirect consequence of training may be that trained participants more frequently self-questioned or reflected on their problem solving strategy. Again, participants trained in inconsistency identification exhibited substantially more of this behavior, although this behavior was less prevalent than the two specifically related training behaviors, discussed above (see Table 1). An ANOVA revealed a significant difference in the amount of self-questioning of strategy, $F(2, 33) = 6.84, MSE = 1.74, p < .01, \eta^2 = .29$, with Bonferroni comparisons confirming that trained participants engaged in significantly more self-questioning than those participants who practiced ($p < .05$) or were not trained ($p < .01$). An example of self-questioning or reflection for the Train problem is, “I don’t really know where to go because I don’t know what the distance is . . . .”

These differences in problem solving, related to the inconsistency identification procedure, provide some validation that the increased solution rate was indeed due to training of this cognitive procedure. Comments from trained participants such as “It doesn’t say in the problem . . . . so it’s not consistent” and “Now I’ll go through the problem again” suggest that participants were check-
ing for inconsistencies between their interpretation and the problem statement. Included in Appendix C is a transcript of a verbal protocol from a trained participant tackling the Light problem, indicating iterative use of the inconsistency identification procedure that eventually overcame the constraint that it was nighttime. Evidence of this in this example comes from 12 rereadings of parts of the problem statement (coded R), eight statements explicitly questioning something in the problem statement (Q), and four instances of self-questioning or reflecting on strategy (SQ). This participant was clearly influenced by the cognitive procedure in training and uses the word consistent on three occasions in this transcript. It is also of interest that the first three hypotheses were innovative and not inconsistent with the problem statement even though they were not correct (see Appendix C).

Training focused on developing a general comparative and evaluative process that had the advantage of being relevant to a variety of verbal insight problems even though how it was instantiated would vary with different individuals and problem constraints. This novel cognitive procedure therefore avoided the criticism that training cued the nature of representation change required by providing information relating to either the specific problem or category of problem. The present study, although capitalizing on the re-encoding mechanism proposed by Ohlsson (1992) in RCT, elaborates how the trained cognitive inconsistency identification procedure can provide an alternative method to the process of representation change based on negative feedback leading to a redistribution of activation (Ohlsson, 2011).

Our conclusions from Experiment 1 should be tempered by two limitations. First, there were two important parts of our training, namely, developing declarative knowledge that made participants aware of the nature of a verbal insight problem (Part 1 of training, Appendix A) and, second, developing procedural knowledge through various forms of practice aimed at effecting the process of representation change (Part 2 of training). The results of Experiment 1 do not allow us to disentangle these effects and identify whether the awareness training alone was sufficient to facilitate representation change or whether practice in inconsistency identification training was necessary. Consequently, we investigated this issue in Experiment 2. Second, it is important to determine over what periods of time the training effect would persist given that, in Experiment 1, testing followed immediately after the training, thus maximizing the opportunity for training to demonstrate a positive effect. Therefore, Experiment 3 was designed to assess the effectiveness of the inconsistency identification training up to 48 hr after it had taken place.

### Experiment 2

The aim of Experiment 2 was to investigate whether the improvement in solution rate was due to the raising of general awareness concerning the nature of verbal insight problems (Part 1 of training, Appendix A) or whether explanation and practice with the inconsistency identification procedure (Part 2 of the training, Appendix A) was also a necessary part of the training.

Surprisingly, the effectiveness of general awareness training on insight problem solving has not been directly investigated. Various studies have provided such awareness within their training, but it has been accompanied with some form of practice so that it is not possible to isolate its effect. Wicker et al. (1978, Experiment 2) used a training condition that instructed participants to reformulate their problem solving strategy in the face of an impasse and not to make unnecessary assumptions. Although this condition was associated with the best solution rate, it also involved considerable time in applying this strategy to eight training problems. Similarly Cunningham and MacGregor (2008) found that training with strategic instructions, warning against making misleading assumptions, produced superior performance to a no-training control condition, although training involved practice with three matchstick problems. In the context of visuospatial insight problems, Walinga, Cunningham, and MacGregor (2011) found the best solution rate was associated with a training condition that focused on mistaken assumptions and barriers in insight problem solving, although avoiding these issues was also practiced during training using the nine-dot problem. To the best of our knowledge, training studies have not investigated whether declarative knowledge that raises general awareness concerning the problematic nature of insight problem solving can be spontaneously translated into an effective strategy without some form of practice, which was therefore the aim of Experiment 2.
There are both empirical and theoretical grounds for predicting general awareness training not to be effective at promoting improved insight problem solving. First, studies using hints either before or during insight problem solving have revealed that even raising awareness through problem-specific hints only has modest success (Burnham & Davis, 1969; Weisberg & Alba, 1981). Extrapolating from these results, one would therefore expect nonspecific instructions to be even less effective. Second, there is strong evidence from a variety of domains that in order for training to effect positive transfer, it is critical that training involves practicing the same cognitive activities that are required by the test or transfer problems. Studies by Lockhart, Lamon, and Gick (1988) and Adams et al. (1988) both stressed the importance of training participants to process information in a problemlike manner, and this concurs with the findings of Needham and Begg (1991) that trainees need to know how to apply relevant information by engaging in transfer-appropriate processing. More recently, a variant of this adage has been termed the “procedural reinstatement principle” that maintains an individual will show high durability and positive transfer performance only when the mental procedures developed during training can be reinstated at the time of testing (Clawson, Healy, Ericsson, & Bourne, 2001; Healy, Wohldmann, Sutton, & Bourne, 2006). This transfer principle is also consistent with the theory of identical elements, originally proposed by Woodworth and Thorndike (1901) with subsequent modifications by Osgood (1949), Gick and Holyoak (1987) and Anderson (1987), emphasizing that positive transfer will increase the greater the similarity between the tasks and skills involved in the training and transfer situations. Therefore, we predict that being made aware of the general nature of insight problems will not improve solution rate without the inconsistency identification training.

Method

Participants. Forty-six psychology students from Cardiff University took part in this study. Ages ranged between 18 and 27 years (M = 19.59, SD = 1.98). Participants were randomly allocated to one of three conditions: the complete inconsistency identification training package (n = 14) (as in Experiment 1), awareness-only training (n = 16), and a no-training control condition (n = 16).

Training and test problems. The same training (Appendix A) and test problems (Appendix B) were used as in Experiment 1 except participants in the awareness training only received the first part of the training that was intended to develop declarative knowledge and awareness that an incorrect interpretation of a problem may block its solution (Part 1 of Appendix A). Participants in this awareness condition then read through two examples of how this blocking effect might operate in two examples, the Push That Car and the Coins of the Realm. Participants in the inconsistency identification training condition received in addition explanation and practice at using the inconsistency identification procedure (Part 2 of Appendix A).

Design and procedure. Participants were individually tested and/or trained. Participants in the no-training condition completed the word association task (Cohen, 1975) as a filler task before being tested. Participants in all conditions attempted the seven test problems that were presented in random order in a text booklet for each participant with the same time limit as in Experiment 1 of 4 min for each problem. Participants wrote down their proposed solution for each problem in the booklet and then rated their familiarity with that problem on a 5-point scale. The total durations of the complete training, the awareness-only training, and the word association control task were approximately 55, 40, and 40 min, respectively.

Analysis of proposed solutions and intercoder reliability.
Five proposed solutions were excluded because the rated familiarity was 4 or 5 on the rating scale. This left 317 proposed solutions. Two coders, the third and fourth authors, independently categorized each proposed solution as correct or incorrect. There was initial agreement on 311 of these, and agreement was reached on the remaining six after further discussion. The Perreault and Leigh index was 0.98, which was satisfactorily high.

Results and Discussion

Participants who received the complete inconsistency identification training (Parts 1 and 2, Appendix A) produced the highest solution rate of 60% compared with 36% solution rates of those in both the awareness training (Part 1 only of Appendix A) and no-training conditions (see Figure 2). An ANOVA conducted on solution rate revealed a significant difference among conditions, $F(2, 43) = 6.12$, $\text{MSE} = .05$, $p < .01$, $\eta^2 = .44$. Bonferroni comparisons revealed that participants who received the complete training package solved more problems than those in both awareness training and control conditions ($p < .01$ and $p < .05$, respectively), and there was no difference in the solution rate of participants in these two latter conditions ($p > .05$).

The results from Experiment 2 demonstrate that providing participants with declarative knowledge concerning the general nature of verbal insight problems together with some examples was insufficient to improve solution rate compared with a no-training condition.
control condition. The higher solution rate for participants who received the full inconsistency identification training replicates the finding of Experiment 1 and demonstrates that practice in trying to spot an inconsistency between one’s interpretation of the problem and the problem statement is a necessary and effective cognitive procedure for improving solution rate.

Experiment 3

To be of practical use, the inconsistency identification training has to be capable of sustaining improved solution rates beyond the immediate posttraining environment. This is particularly important as training has the challenging task of mitigating the powerful effect of the wrong stereotypical representation being generated after reading the problem statement. In terms of Ohlsson’s (2011) activation account, we need to ascertain that there is sufficient activation strength for the cognitive procedure to be recalled and applied to posttraining problems at least up to 48 hr after training.

There have been many reviews of the retention of trained skills that have been driven by the need to minimize any deterioration in performance level over time, particularly within safety-critical environments (e.g., Hagman, 1983; Healy & Sinclair, 1996; Naylor & Briggs, 1961). Also Arthur, Bennett, Stanush, and McNelly (1998) carried out a meta-analysis of data from 53 articles to estimate the size of effect of different factors. These reviews agree that aspects of the task, the design of training, the nature and duration of the retention period, and the transfer situation all influence the retention of skilled behavior. Factors include the nature and extent of the skill to be learned, the degree of original learning, the duration and conditions of the retention interval, and the relationship between the training and transfer situations. It is not our intention, nor is it possible to calibrate our inconsistency identification training procedure on all of these dimensions and then predict systematically the extent of recall after a delay period with no practice. However, we highlight below some important considerations.

First, the nature of the inconsistency identification procedure can be reduced to a couple of high-level rules concerned with trying to spot an inconsistency between the problem specification and one’s interpretation of it, and iteratively selecting and considering different parts of the problem specification. This procedure is relatively brief compared with the procedural tasks studied in the above retention literatures that may involve many discrete steps. In addition, these two higher level rules are integrated and organized into one overall procedure, a feature that facilitates retention (Naylor & Briggs, 1961). At the same time, the cognitive procedure is unique in both its nature and the situation in which it is to be applied and therefore is unlikely to suffer interference from other skills during the training, retention, or testing periods. Also, the training is focused on transfer-appropriate processing, as discussed above, and trainees practice three example problems that require the same cognitive activities as the test problems. In addition, the training and testing contexts involve the same laboratory with the same experimenter. All of these features led us to consider that there will be little if any reduction in performance over a couple of days after training.

Results and Discussion

Participants who received inconsistency identification training produced higher solution rates than those in the no-training control condition irrespective of whether training was tested immediately after training or was delayed by 24 or 48 hr (see Figure 3). Solution rates of 65%, 57%, and 62% were achieved by participants in the immediate, 24-, and 48-hr-delayed test conditions, respectively, in comparison with a 25% solution rate of participants in the control condition. The higher solution rate for participants who received the full inconsistency identification training replicates the finding of Experiment 1 and demonstrates that practice in trying to spot an inconsistency between one’s interpretation of the problem and the problem statement is a necessary and effective cognitive procedure for improving solution rate.

Method

Participants. Fifty-eight psychology students from Cardiff University took part in this study. Ages ranged between 18 and 47 years \( (M = 19.72, SD = 3.88) \). Participants were randomly allocated to one of four conditions: the full inconsistency identification training package (as in Experiments 1 and 2) with testing immediately after training \( (n = 13) \), or delayed by 24 \( (n = 15) \) or 48 hr \( (n = 15) \), and a no-training control condition \( (n = 15) \).

Training and test problems. The same training and test problems were used as in Experiment 1 (see Appendices A and B).

Design and procedure. Participants in the no-training control condition completed the word association task (Cohen, 1975) as a filler task before being tested. Other features were the same as Experiment 2. The total duration of each training condition was approximately 55 min that was split approximately equally between training and testing for the two delayed testing conditions. The duration of the word association control task condition was approximately 45 min.

Analysis of proposed solutions and intercoder reliability.

Ten proposed solutions were excluded because the rated familiarity was 4 or 5 on the rating scale. This left 396 proposed solutions. Two coders, the third and fourth authors, independently categorized each proposed solution as correct or incorrect. There was initial agreement on 388 of these, and agreement was reached on the remaining eight after further discussion. The Perreault and Leigh index was .98, which was satisfactorily high.

Results and Discussion

Participants who received inconsistency identification training produced higher solution rates than those in the no-training control condition irrespective of whether training was tested immediately after training or was delayed by 24 or 48 hr (see Figure 3). Solution rates of 65%, 57%, and 62% were achieved by participants in the immediate, 24-, and 48-hr-delayed test conditions, respectively, in comparison with a 25% solution rate of participants in the control condition.
condition. An ANOVA conducted on solution rate revealed a significant difference among conditions, $F(3, 54) = 8.84, MSE = .05, p < .001, \eta^2 = .44$. Bonferroni comparisons revealed that participants trained in the immediate, 24-, and 48-hr-delayed testing conditions solved more problems than those in the no-training control condition ($p < .001$, $p < .01$, and $p < .001$, respectively), and there was no significant difference in the solution rate of participants in these three training conditions ($ps > .05$).

The results from Experiment 3 are straightforward in demonstrating that the cognitive procedure involved in training could be recalled and applied successfully to solve verbal insight problems as long as 48 hr after initial training. This is important theoretically, as the trained procedure retained sufficient activation to be recalled despite the strong stereotypical tendency to assume that the default representation was correct. The result is also important from a practical perspective in demonstrating that training did have a positive transfer effect on problems encountered up to at least 2 days later.

**General Discussion**

We have proposed in the present study the use of a novel cognitive procedure to effect representation change in verbal insight problems. The technique assumes an incorrect interpretation will be initially developed on reading the problem, and this situation provides the opportunity for problem solvers to attempt to identify some inconsistency between the problem specification and their initial interpretation of the problem. Identification of such an inconsistency not only acts as a cue to change a wrong interpretation of the problem or reencode the problem in Ohlsson’s terms, but the nature of the inconsistency may also suggest how re-encoding should take place. This training technique therefore involves a generic cognitive procedure for solving verbal insight problems specified in this form that does not provide any problem-specific cue(s) concerning, for example, details of the problem or category of problem.

This inconsistency identification training does not guarantee solution, although there is strong and consistent evidence from Experiments 1–3 that it substantially improved solution rates between 24% and 40% against a no-training control condition. Solution rates for this training ranged between 60% and 68% when tested immediately after training, and the results from Experiment 3 revealed that there was no decrement in performance if testing was delayed for 24 or 48 hr. Furthermore, we are able to reject the possibility that improved problem solving was due to raised awareness concerning the nature of verbal insight problem solving and the need to reformulate one’s thinking rather than explanation and practice in using the inconsistency identification procedure. The results from Experiment 2 indicate that providing only such general awareness in training does not improve solution rate.

Experiment 1 is particularly important as we gathered reliable supplementary evidence, from an analysis of verbal protocols, of how thinking changed in line with the trained cognitive procedure. Such an analysis has been suggested by Ash et al. (2009) as necessary in order to validate that training does affect cognitive strategy in the manner predicted and therefore is responsible for any improvement in performance. Without such data, there is no direct evidence of how training affected the process rather than the product of problem solving. Given the nature of the inconsistency identification training, we were initially surprised that participants in this condition in Experiment 1 did not reduce the number of inconsistent hypotheses explicitly generated. However, it should be noted that inhibiting the verbalization of inconsistent hypotheses was not directly trained, unlike the encouragement to engage in rereading and questioning behaviors that directly related to the trained procedure. There was substantial evidence that training did affect the nature of problem solving behaviors in the manner intended by increasing rereading or paraphrasing of the problem statement and explicit questioning of detail in the problem statement. Consequently, this evidence provides some validation that training did achieve its goal in how thinking was changed, and therefore training was responsible for the higher solution rate. Besides these specific behavior changes directly related to training, there was also some modest evidence that inconsistency identification training also affected higher level metacognitive processes. Participants who received this training engaged in more evaluation of and reflection on their problem solving strategy. This was an indirect effect of the training regime. The association of higher solution rates with an increase in self-questioning of both the problem details at the lower level and strategy at the higher level is consistent with positive evidence from training studies that have manipulated self-questioning, sometimes under the banner of metacognitive techniques (e.g., Ainsworth & Burcham, 2007; Tajika, Nakatsu, Nozaki, Neumann, & Maruno, 2007). Tajika et al. (2007) suggested that certain metacognitive instructions (self-questioning and self-examination) may act to determine problem misunderstandings and generate more relevant problem information. In a similar research vein, instructing participants to explain or justify their solutions has been demonstrated to aid performance in problem solving (Ahluum-Heath & DiVesta, 1986; Berry & Broadbent, 1984, 1987) and design tasks (Wetzstein & Hacker, 2004; Winkelmann & Hacker, 2010). Although our results are consistent with these research findings, it should be noted that the inconsistency identification training focused on a specific evaluative process concerning extra or incorrect information that the problem solver assumed wrongly was part of the problem specification.

We need to acknowledge various shortcomings of the present article. First, the inconsistency identification training is designed to facilitate the solution of verbal insight problems that have a particular structure, namely, a written specification that is incomplete and contains familiar details likely to be misinterpreted by default reasoning. The majority of verbal insight problems that have been studied have these features and are therefore in scope to improvement from our trained cognitive procedure. However, Remote Associate Test problems and rebus puzzles have been discussed in relation to insight (Cunningham, MacGregor, Gibb, & Haar, 2009), although their structure would not be compatible with the application of our procedure. Second, the design and rationale of the cognitive procedure has been discussed within the context of RCT, which discusses mechanisms for representation change. However, the inconsistency identification procedure could also be conceptualized within the criterion for satisfactory progress theory (MacGregor, Ormerod, & Chronicle, 2001) as another mechanism for triggering change of the problem space when satisfactory progress toward the goal can no longer be achieved.

In conclusion, developing representational change to gain insight is important and challenging because the incorrect assump-
tions that block problem solution concern well-learned, high-frequency responses that are implicit and automatic and would be correct on the majority of occasions in other contexts. We have developed a novel cognitive procedure for operationalizing how the re-encoding, suggested by RCT (Ohlsson, 1992, 2011), can be cued more directly rather than waiting for the gradual accumulation of negative feedback to result in a deactivation of the inappropriate representation. This involves training problem solvers to use a comparative evaluative process that attempts to identify any inconsistency or unwarranted assumption between the problem solver’s interpretation of the problem and the problem statement itself. Unlike other studies (e.g., Patrick & Ahmed, 2014), the inconsistency identification procedure does not use detail that is specific to one problem or a category of problem and therefore is potentially applicable to many verbal insight problems with different constraints.

References


Appendix A

Inconsistency Identification Training Program

Training comprised two parts and was presented as typed text in a paper booklet. The aim of the first part was to provide general awareness of the nature of verbal insight problems and how an incorrect interpretation can block solution. The second part of training concerned using an iterative cognitive procedure that aimed to identify any inconsistency between a person’s interpretation of any part of the problem and the problem specification.

**Part 1**

The training material was as follows:

A common feature of verbal insight problems is that a person’s usual interpretation will contain some detail or unwarranted assumption that is inconsistent with the problem statement. An incorrect interpretation of a problem may then block its solution.

For instance, consider the following problem:

A man pushed his car. He stopped when he reached a hotel at which point he knew he was bankrupt. Why?

Often, an inappropriate interpretation is activated on the basis of past experience such that, in the absence of further information, the car is assumed to be a real car. The car is in fact a car on a monopoly board.

Now consider another problem:

Why are 1988 pennies worth more than 1983 pennies?

It is often assumed that the numbers refer to dates, although this is not actually specified in the problem statement. They actually refer to quantities, and hence 1,988 pennies are worth £19.88, whereas 1,983 pennies are worth £19.83.

**Part 2**

Subsequently, participants were told that they would learn a simple procedure that would help them solve these problems and they would have practice in using it. Participants were instructed in text that they should use the following procedure in order to try and identify any inconsistency between their interpretation of any part of the problem and the problem specification.

After reading the problem carefully, select part of the problem and consider your interpretation of it. Then check whether your interpretation is inconsistent by comparing it with the information in the problem specification. If it is inconsistent, then consider how you can make it consistent and whether your revised interpretation can be used to solve the problem. If not, then select another part of the problem and repeat the above process until you have generated a solution to solve the problem.

Next participants read the Sid Shady problem in the text booklet, and then it was explained how three interpretations of the problem situation were inconsistent with the problem statement, as follows:

To answer the question, a possible interpretation concerning what Shady was stealing is that something was concealed on his person. However, this is inconsistent with the problem statement as it does not explain the taking of the lumber, wires and concrete. Another interpretation is that Shady was using the lumber, wires and concrete for a useful purpose but this is also inconsistent as they are of no value and the guard was aware that he was taking them. An alternative interpretation is that Shady was concealing something within the contents of the wheelbarrow, which is also inconsistent because the statement states that the wheelbarrow is checked daily. The correct interpretation and solution is that Shady is stealing wheelbarrows, which is consistent because the statement states that Shady leaves with a wheelbarrow each night.

Finally, training provided participants with practice at identifying how their stereotypical interpretations of two verbal insight problems were inconsistent with the problem statements. Both problems were again presented by typed text, and solution feedback was provided by the experimenter after each problem. Participants practiced using the inconsistency identification procedure with the Barney Dribble problem and, if they failed to do so, were prompted by the experimenter. Finally, participants attempted to solve the Coffee problem without any support, and only solution feedback was provided.
Appendix B

Training and Test Problems

Training Problems

PUSH THAT CAR
A man pushed his car. He stopped when he reached a hotel at which point he knew he was bankrupt. Why? (Sloane, 1992)

Constraint: Car refers to a real car.

COINS OF THE REALM
Why are 1988 pennies worth more than 1983 pennies? (Sloane, 1992)

Constraint: 1988 and 1983 refer to dates.

SID SHADY
Sid Shady works for a large construction company that was very concerned about employee theft. Someone tipped off the company that Shady was the man to watch. Each night he passed through security with a wheelbarrow full of scrap lumber, discarded electrical wires and chunks of concrete. The security guards checked the contents daily but could find nothing of value. What was Shady stealing? (Ansburg & Dominowski, 2000).

Constraint: What was stolen was within the wheelbarrow.

BARNEY DRIBBLE
Barney Dribble is carrying a pillow case full of feathers. Hardy Pyle is carrying three pillow cases the same size as Barney’s, yet Hardy’s load is lighter. How can this be?

Constraint: Hardy’s pillow cases were full.

COFFEE
A woman said to her husband “This morning, one of my earrings fell into my coffee. Even though my cup was full, the earring did not get wet.” How could this be true?

Constraint: Coffee was liquid coffee.

Test Problems

PEAR TREE
A farmer in California owns a beautiful pear tree. He supplies the fruit to a nearby grocery store. The store owner has called the farmer to see how much fruit is available for him to purchase. The farmer knows that the main trunk has 24 branches. Each branch has exactly 6 twigs. Since each twig bears one piece of fruit, how many plums will the farmer be able to deliver? (Ansburg & Dominowski, 2000).

Constraint: A calculation is needed and the same fruit is involved.

DR APPLE
Shadow opened the door to Dr. Apple’s office and surveyed the scene. Dr. Apple’s head lay on his desk in a pool of blood. On the floor to his right lay a gun. There were powder burns on his right temple indicating that he was shot at close range. On his desk was a suicide note, and in his right hand was the pen that had written it. Shadow noted that death had occurred in the last hour. All of a sudden Dr. Apple’s wife burst into the office and screamed “My husband’s been shot!” She ran toward the body and saw the note and cried, “Why would he want to kill himself?” Shadow replied “This was no suicide; it is a clear case of murder.” How does Shadow know? (Ansburg & Dominowski, 2000).

Constraint: Sequence of occurrence of events and that the pen could have been in Dr Apple’s hand if he had committed suicide.

TRAIN
At 7 a.m., a train moving 90 mph leaves Montreal heading for Toronto. At 8 a.m., a train running 110 mph leaves Toronto heading for Montreal. Which train will be closer to Montreal when they meet?

Constraint: A calculation is needed.

DIRECTORY
There is a town in Northern Ontario where 5% of all the people living in the town have unlisted phone numbers. If you selected 100 names at random from the town’s phone directory, on average, how many of these people selected would have unlisted phone numbers? (Sloane, 1992)

Constraint: That a phone directory contains unlisted numbers.

ANTIQUE COIN
A dealer in antique coins got an offer to buy a beautiful bronze coin. The coin had an emperor’s head on one side and the date 544 BC stamped on the other. The dealer examined the coin and realized it was a fake. How did he know the coin was phoney? (Ansburg & Dominowski, 2000).

Constraint: The date on the coin is correct.

PROFESSOR BUMBLE
Professor Bumble, who is getting on in years, is growing absent-minded. On the way to a lecture one day, he went through a red light and turned down a one-way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behaviour?

Constraint: Professor Bumble was driving.

LIGHT
A young boy turned off the light in his bedroom and managed to get into bed before the room was dark. If the bed is ten feet from the light switch and he used no wires, strings or any other contraptions to turn off the light, how did he do it?

Constraint: It was dark.
Appendix C

An Example of a Trained Participant’s Coded Verbal Protocol When Tackling the Light Problem (Experiment 1)

| R1, R2  | “Um...<turned off the light in his bedroom > ...and <managed to get into bed before the room was dark>. <I’m not really sure if the room is his bedroom>, <H1, so he could possibly be getting into bed in another room>... <If the bed is ten feet from the light switch>...<I’m not ever so sure that’s got anything to do with it at the minute>...<A young boy turned off the light in his bedroom and managed to get into bed before the room was dark>... <H2, possibly there was a light source elsewhere>, outside...<Got into bed before it was dark>...<H3, The light may have possibly had a setting on it that stayed on...for a few seconds after>...<The bed is ten feet from the light switch>. I’m not sure what or how that ten feet will be-- <maybe he could turn it off from there> I’m not sure, whilst he’s in bed.<cl it really consistent with the problem?> It is possible that the light has some sort of setting if it is the main light. <I’m assuming that it was the main one>...Maybe he went...oh no <there were no other contraptions>...so that <the possibility of turning it off from the bed probably isn’t consistent with the problem>...<Turned off the light in his bedroom>. <At the moment I’m thinking>...a solution...probably something to do with the light having a setting...<If the bed is ten feet>, it’s quite far I think isn’t it from the light switch...and he used no wire...<So a young boy, he turned off the light in his bedroom before the>... <unless the bed’s in another room but then>...<he managed to get into bed before the room was dark>...<possibly we shouldn’t assume that it’s the night time>...<possible solution yeah it might be that. He may have turned off the light in his bedroom and it wouldn’t have made any difference>, <H4, if it was daytime outside> and therefore he managed to get into bed before the room was dark. <That’s possibly consistent with the problem> <so if there was no wires, strings or contraptions>...and he can’t reach the light switch...That may be...<a young boy turned off the light in his bedroom and managed to get into bed before the room was dark>...<I think the answer is that the boy did turn off the light in his bedroom and got into bed before the room was dark after it was still sort of daytime and there was light coming possibly from a window.”

Note: Coding categories: R = any time a participant rereads or paraphrases whole or part of the problem statement during problem solving; Q = a statement explicitly questioning something in the problem statement; SQ = self-questioning or evidence about a reflection concerning participants own explanations, inferences, or strategy (i.e., reflecting on their own thinking). Bold font prefaced by H = a proposed hypothesis. Each verbal item relating to a category is enclosed by <<>>. The inclusion of numbers denotes the order of occurrence of the verbal items within a category.

Received August 20, 2013
Revision received May 29, 2014
Accepted June 3, 2014