Running Head: NEGATIVE EFFECTS OF EVALUATION

Explaining Away the Negative Effects of Evaluation on Analogical Transfer: The perils of premature evaluation.

Christopher Bearman

University of South Australia, Australia

Thomas. C. Ormerod*

Lancaster University, UK

Linden. J. Ball

Lancaster University, UK

Daniel Deptula

University of South Australia, Australia

*Corresponding Author:

Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK

Tel: +44 (0)1524 593164

Fax: +44 (0)1524 593744

Email: t.ormerod@lancaster.ac.uk

Explaining Away the Negative Effects of Evaluation on Analogical Transfer: The perils of premature evaluation

Abstract

Four experiments explored effects on analogical transfer of evaluating solutions to base problems. In contrast to reports of positive effects of explanation, evaluation consistently reduced transfer rates and impaired mental representations of base material. This effect was not ameliorated by encoding for a later memory test, summarising, or engaging in similar processes at encoding and recall. However, providing a prior explanation task removed the inhibitory effect of evaluation. It appears that evaluation leads to encoding of extraneous material that interferes with access to solution-critical analogous information. Prior explanation inoculates against negative effects on transfer by ensuring that new information introduced via evaluation is organised around existing representations of relevant information of the base problem. The results suggest that the source of difficulty in analogical transfer may reside not only in retrieval and mapping but also in the initial encoding of problems.

Keywords: Analogical transfer; Problem solving; Evaluative processing; Self-explanation; Encoding.

Explaining Away the Negative Effects of Evaluation on Analogical Transfer

Analogical problem-solving involves the transfer of information from one context to another conceptually-related context. Such transfer has often been difficult to achieve in laboratory studies (Anoli, Antonietti, Crisafulli, & Cantoia, 2001; Gick & Holyoak, 1980, 1983). For example, Gick and Holyoak (1980, 1983) found that providing a base analogue produced only modest increases over baseline solution rates where no base analogue had been provided (20-30%).

Several manipulations have been shown to improve transfer of the base analogue without a hint, such as providing participants with an animated diagram (Pedone, Hummel, & Holyoak, 2001), increasing the surface similarity of the base material to the target problem (Holyoak & Koh, 1987) and encouraging participants to compare two base analogues (Catrambone & Holyoak, 1989; Gentner, Loewenstein, & Thompson, 2003; Gick & Holyoak, 1983). Encouraging participants to compare two base analogues seems to improve transfer because it facilitates encoding of the base material by focusing learners on its relevant aspects (Catrambone & Holyoak, 1989; Gentner et al., 2003). To put this another way, comparing two base analogues makes the relevant aspects more salient in the mental representation that is formed of the base material.

The way in which participants represent the base material therefore appears to be crucial in determining whether the base material will transfer to a conceptually related problem. Base information that can be easily integrated into a participant's existing knowledge, such as concepts that are well known to the participant (Schunn & Dunbar, 1996) or concepts that are located in a domain for which the participant has high aptitude (Blessing & Ross, 1996; Novick, 1988) usually show good rates of transfer. In addition, manipulations such as *explanation,* that seek to integrate base material with existing knowledge also improve transfer. Explaining material to oneself or others has been found to have beneficial effects on the later recall and use of that information (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, DeLeeuw, Chiu, & LaVancher, 1994; Neuman & Schwarz, 1998). Chi et al. (1989) found that participants who engaged in self-explanation while attempting to understand information scored better on a subsequent knowledge test than participants who simply talked aloud. Chi et al. (1994) further demonstrated that instructing participants to self-explain led to improved performance on isomorphic problems compared to participants who were asked simply to read aloud. Explanation has also been found to reduce errors, particularly those errors that are related to gaps in knowledge (VanLehn & Jones, 1993). Chi (2000) argues that attempting to understand information through self-explanation leads to the formation of inferences about that information in response to perceived gaps in knowledge. This generation of inferences leads to the revision or re-organisation of the person's knowledge structure, which creates a richer, more elaborated understanding of the information (Pirolli & Anderson, 1985).

A mental process that appears to have similar characteristics to explanation is *evaluation.* Evaluation is defined here as the judgement of problem-related information under qualitative valances such as importance, value, appropriateness and scope (Bearman, Ormerod, & Ball, 2003). Evaluation is required when solutions to problems have been produced and is an important part of making decisions in real world domains (Easton, 1992; Darses, 2002; Klein, 1998). Evaluation has been seen as an important part of human intellectual performance, forming a central part of Miller, Galanter, and Pribram's (1960) theory of cognition. Miller et al. proposed the test-operate-test-exit (TOTE) model as a basic building block of cognition, thus allocating a central role to the test (or evaluate) component. Moreover, encouraging evaluation appears to reduce the tendency of auditors to over-rely on recent cases when making decisions (Ashton & Kennedy, 2002).

Like explanation, evaluation is a constructive activity that builds knowledge. Evaluation requires that participants examine each piece of information in a detailed way, drawing on prior knowledge to assess its value. This process allows the integration of new information with existing knowledge structures. During this process, gaps in the participant's mental model of the situation should be identified and inferences made that fill those gaps, thus creating a more comprehensive mental model of the information. We might, therefore, expect evaluation to show the same kinds of facilitation on analogical transfer as explanation.

A number of studies have observed a *decrement* in performance using a manipulation referred to as *justification* (e.g., Sieck, Quinn, & Schooler, 1999). Sieck et al. found that providing justification of comparison decisions increased the use of superficial information at the expense of structural information. Sieck et al. theorise that this performance decrement is due to hard-to-verbalise information being neglected in the participant's mental representation relative to other more easily verbalised information. However, the concept of justification used in these studies refers to the provision of an explanation of one's *decisions* rather than an explanation of a *problem and solution*. Likewise, it differs from evaluation in focussing on participants' meta-cognitions of their own thinking rather than evaluations of the base problem. In the present studies there is no reason to assume *a priori* that there is a difference between evaluation, explanation and processes such as summarisation in ease of verbalisation, since all of these tasks focus upon the content of the base problem. Hence, it is expected that evaluation will facilitate analogical transfer. The exploration of this hypothesis forms the basis for the work reported here. We report four experiments that explore the effects of evaluation on analogical transfer.

Experiment 1a

Experiment 1a investigated the effects of instructions to *evaluate* upon analogical problem solving using similar materials to those of Gick and Holyoak (1980). If evaluation acts as a form of explanation, then instructions to evaluate base material should enhance performance on the later test problem, thus demonstrating enhanced analogical transfer of the base material. Also, since the participants' encoding of the base material seems to be crucial for transfer, half of the participants were told that they would receive a memory test on the base information (that was not actually given). This was designed to make the information more salient in the mental representation that participants formed of the base stories. Instructing participants to memorize base material has been found previously to increase solution rates to the target problem (Needham & Begg, 1991). Instructions to memorize base information might be additive or interactive with the effects of instructions to evaluate.

Method

Design. A two-factor, between-participants design was used, with base processing (evaluative vs. non-evaluative) as one factor, and memory instruction (memory encoding vs. no memory encoding) as the other factor. The dependent variable was the production of the "dispersion-convergence" solution (presence or absence), assessed pre- and post-hint.

Participants. 37 females and 35 males participated in the experiment; all bar 12 were students attending Lancaster University. Participants were paid £3 for the experiment, which took approximately 30 minutes. Participants were randomly allocated to the different conditions.

Materials. The materials were similar to those of Gick and Holyoak (1980, 1983), with the target task being Dunker's (1945) "radiation problem" (see Appendix). Since the base story ("The General") used by Gick and Holyoak tends to produce ceiling effects on post-hint performance (i.e., after participants are told to use the base problem as an analogue to solve the target problem), a new base was developed to convey the critical dispersion-convergence solution principle. This base involved a story about a nuclear power station that required a constant supply of water in order to operate safely. However, one single pipe would not be sufficient to produce this constant supply and the solution was to build multiple pipes to carry the required volume (see Appendix). This base analogue could be used by participants to generate solutions to the radiation problem at either a conceptual level (replacing a large force with multiple smaller forces) or at a superficial level (invoking surface-level entities such as pipes). The nuclear power station story does not include any reference to a central location or to multiple directions (as Gick & Holyoak's The General story does). These components are critical for the correct solution to the target problem and so the nuclear power station story is less adaptable to the target problem than is Gick and Holyoak's The General story (Keane, 1996), but their presence in the latter might be seen to make the analogous components overly salient and therefore trivial to map as literal similarities (cf. Gentner, 1983) onto a target problem. A pilot study of six participants indicated that the nuclear power station story produced intermediate levels of transfer (50% post-hint solution rates compared to 92% in Gick & Holyoak, 1980).

The nuclear power station story presents a base that is superficially different but conceptually similar to the radiation problem. Two distracter stories were also developed. The first was superficially similar but conceptually different to the radiation problem, and referred to an osteopath who wished to cure back pain using a novel machine (see Appendix). The second distracter was a *non-analogous* story that was neither conceptually similar nor superficially similar to the radiation problem, and described a commander overthrowing a military headquarters using a big bridge. The non-analogous story was loosely based on Gick and Holyoak's (1983) "The Commander" story (see Appendix).

Procedure. Although the basic experimental procedure closely followed the one employed by Gick and Holyoak (1980, Experiment 4), a subtle procedural change was instigated to encourage participants in the evaluation conditions to believe that each presented solution was not necessarily the most effective one for solving a particular base problem, thereby promoting enhanced evaluative processing of that solution. Participants were asked to read through the base problem, and were then asked to select at random one from a set of five cards that actually contained identical solutions to the base problems. Participants were then required either to "evaluate whether the suggestion is a good solution to the problem" (evaluative instructions) or to "read out loud and summarize the problem and suggested solution" (non-evaluative instructions). Participants in the memory conditions were instructed to attempt to memorize base problems and solutions for a later recall test (which was not actually given). Each base story was processed for two minutes. If participants lapsed into silence for more than five seconds they were asked to keep talking. If participants still did not talk they were asked to continue according to their condition (i.e., "please keep evaluating" in the evaluation condition). The order of base stories was counterbalanced across participants. Participants then worked on Wason's (1960) 2-4-6 task for ten minutes. Finally, participants attempted the radiation problem. If the dispersion-convergence solution was not given within five minutes, a hint was presented that "one or more of the stories seen at the beginning of the experiment could be used to form a solution to the [target] problem", and participants were given a further three minutes to solve the problem. Participants were scored as having produced the dispersion-convergence solution when they included both low intensity and multiple x-rays and implied that these are to be applied simultaneously. This scoring procedure rules out the solution of giving multiple small doses applied separately over time, and is a stricter criterion of success than that used by Gick and Holyoak.

Results and Discussion

Six participants generated the dispersion-convergence solution to the radiation problem without a hint to use the previous stories. Of these participants, one was in the memory/evaluation condition, two were in the no memory/no evaluation condition and three were in the memory/no evaluation condition. This represents a rate of 8% dispersion-convergence production across groups. This is comparable to the rates observed by Gick and Holyoak (1980) for an analogous story and two distracters. The post-hint frequency of dispersion-convergence production is presented in Table 1.

(Insert Table 1 about here)

A logistic regression using base processing and memory instruction as predictors yielded a significant model, $\chi 2(1, N = 72) = 7.7$, $p = .02$, with base processing the only significant predictor in the model, Wald = 6.70, $p = .009$. Table 1 shows that instructions to evaluate have a detrimental effect on later analogical transfer.

In contrast to the prediction that evaluative base processing would lead to improved analogical problem solving, evaluating the base information made participants significantly less likely to produce the analogous solution to a target problem. One possible explanation for this negative effect is that participants did not spend as long on encoding the base problem when they evaluated the presented solution compared with the control group. However, if it were simply an effect of processing time, one would expect the effect to be ameliorated by instructing participants to remember the story with the view to later recalling it. Yet, participants who were asked to memorize the story having evaluated the solution showed no advantage over those who evaluated without memorization. Thus, the effect of evaluation seems to be genuinely inhibitory and not just an artefact of processing time or effort. In light of the unexpected nature of the findings, Experiment 1b was conducted to determine whether this effect is robust.

Experiment 1b

In addition to exploring the replicability of a negative effect of evaluation, Experiment 1b investigated whether instructions aimed at encouraging participants to evaluate their solutions to the *target* problem might reduce the detrimental effect of base evaluation. The rationale behind this manipulation derives from *transfer appropriate processing*. Roediger (1990) found that recall of information is reliably enhanced when the processes engaged in at retrieval match those during encoding. Dunbar, Blanchette, and Chung (cited in Dunbar, 2001) found that highlighting the identification and recall of relational and structural features in both base and target problems increased analogical transfer. Likewise, Needham and Begg (1991) found facilitation for memory tasks with memory encoding of the base story, and facilitation for problem-solving tasks with solution-oriented encoding of the base. The findings of Dunbar et al. and Needham and Begg lend support to the idea that the detrimental effect of evaluation might be overturned if participants engage in evaluative processing of target material. In Experiment 1b, we assessed transfer appropriate processing by crossing the instructions that participants received during base processing (evaluative vs. non-evaluative) with those that they received during target processing (evaluative vs. non-evaluative).

Experiment 1b also included a procedural change designed to investigate participants' representation of the base information. If encoding of the key elements of the base information is impaired in the evaluation conditions then it would be expected that recall of that information would be worse compared to other conditions. Thus, following target problem-solving, participants were asked to recall the analogous base story in as much detail as possible.

Method

Design. There were two between-participants factors: base processing (evaluative vs. non-evaluative) and target processing (evaluative vs. non-evaluative). Dependent variables were production of the dispersion-convergence solution (presence or absence) assessed pre-hint and post-hint, a measure derived from the number of correct solution elements in each participant's best solution, and a measure of recall of the analogous base story.

Materials. The materials were those used in Experiment 1a.

Participants. 39 female and 33 male students from Lancaster University participated in the experiment. Participants were paid £4 for the experiment which took approximately 40 minutes. Participants were randomly allocated to the different conditions.

Procedure. The procedure of Experiment 1a for conditions without a memory instruction was used here. The only difference was that, prior to attempting to solve the target (radiation) problem, participants were instructed, either: (1) "It does not matter if you do not think the solutions will work, as the aim is to brainstorm the problem" (i.e., non-evaluative target processing); or (2) "Please try not to provide solutions that will not work, as the aim is to produce the best possible solution" (i.e., evaluative target processing). In addition, after participants had attempted (whether successfully or unsuccessfully) to solve the radiation problem, they were requested to recall the "meaningful content" of the story about the nuclear power station (the analogous base).

Results

One participant produced the dispersion-convergence solution pre-hint. This participant was in the non-evaluative encoding and non-evaluative target processing condition. Post-hint dispersion-convergence production rates are presented in Table 2.

(Insert Table 2 about here)

A logistic regression using base processing instruction and target processing instruction as predictors of success yielded a significant model, $\chi 2(1, N = 72) = 13.0$, $p = .01$, with base processing instruction being the only significant predictor in the model, Wald = 7.03, $p = .008$.

Participants' solutions to the target problem (the radiation task) were examined for the number of solution components that they contained. In order to be completely correct, participants must mention *lower* forces, *multiple* forces, *converging* forces and *x-rays*. Thus, a four-point score can be derived which assesses the extent to which participants are close to the dispersion-convergence solution. Although this measure will be correlated with task success (in that a person producing a correct solution will have generated all four critical solution components) it provides a more sensitive score that encompasses nearly-correct solutions. Participants who evaluated the base information had a significantly lower mean score for the critical solution components ($M = 1.71$, $SD = 0.82$) than participants who did not evaluate the base information ($M = 2.11$, $SD = 0.73$), $F(1, 65) = 4.52$, $p = .03$.

In order to investigate further what may be causing the detrimental effect of base-evaluation instructions, participants' recall of the analogous power station story was assessed after they had attempted to solve the target problem. The nuclear power station story can be decomposed into several critical features according to what is relevant for later transfer to the radiation problem. These features include *lower* forces in the pipes (corresponding to lower intensity X-rays in the radiation problem), *multiple* pipes (analogous to multiple X-rays), and *converging* forces (the same in both solutions). The presence or absence of these features in the participants' recall of the nuclear power station story thus formed a three-point coding scheme for investigating recall of the critical components of the base story.

Participants who had evaluated the base information produced a significantly lower average score ($M = 1.84$, $SD = 1.62$) than participants who had not evaluated the base information ($M = 2.65$, $SD = 1.66$), $F(1, 65) = 4.42$, $p = .04$. Participants in the evaluative base-processing conditions were less likely than those in the non-evaluative conditions to mention "lower forces" (18 vs. 24) and "converging forces" (9 vs. 14). This suggests that participants are forming an impoverished representation of the information contained in the base.

Discussion

The detrimental effect of evaluative processing of base stories on post-hint analogising found in Experiment 1a was replicated in Experiment 1b. Engaging participants in transfer appropriate processing (i.e., requiring evaluation of the target problem as well as the base information) did not remove this effect. In addition, participants who evaluated the base information did not retain as much of that information as people who merely summarized the material. It therefore seems that focusing processing effort on the quality of a solution leads to important information being lost from the representation of the base analogue.

Experiment 1c

In Experiments 1a and 1b, summarizing the base information was more effective than evaluation. Thus, it may be possible that instructing participants to produce a summary following base processing will ameliorate the negative effects of evaluation. Encouraging participants to form abstract representations of the base information has been found to increase spontaneous transfer to the radiation problem (Catrambone & Holyoak, 1989; Mandler & Orlich, 1993). Mandler and Orlich (1993) found effective transfer when participants produced a summary of the goal, dilemma and solution in the base information; an effect enhanced still further when participants also produced a statement that encapsulated the general solution principle. Experiment 1c examined the effects of producing different types of summary on evaluation in analogical transfer.

Experiment 1c also allowed a test of whether extending the time available for base processing, a necessary outcome of summarization, might enhance analogical transfer rates after evaluation, relative to the shorter base processing times of Experiments 1a and 1b. Extending the time available for base processing examines the notion that the negative effect of evaluation stems from reduced time available for participants to process the base information due to the additional task requirement of deciding on the quality of the solution.

Method

Design. A two-factor, between-participants design was used, with base processing (evaluative vs. non-evaluative) and solution principle (generated vs. absent) as factors. The dependent variables were as in previous experiments.

Participants. 42 female and 30 male students from Lancaster University participated in the experiment. Participants were paid £4 for the experiment, which took approximately 50 minutes. Participants were randomly allocated to the different conditions.

Materials. The materials were the same as those used in Experiment 1a.

Procedure. The same procedure as Experiment 1a was used (without the memorization instructions), except that after participants had processed the base stories (either evaluating or summarizing out loud) they were presented with a sheet that requested them to identify the aims of the main actor(s) in the story, the problem facing the actor(s) and the solution that the actor(s) came up with. Participants in the solution principle condition were also requested to identify a solution principle for the stories, which was further explained as being "the moral of the tale". As in previous experiments participants processed each base according to condition for two minutes. In addition, in this experiment participants spent an additional two and a half minutes on each base story completing the summarization sheet.

Results

Eight participants (11%) produced the dispersion-convergence solution to the radiation problem pre-hint. Four were in the control condition without a solution principle criterion, two were in the control condition with a solution principle criterion and two were in the evaluation with solution principle criterion condition. Post-hint dispersion-convergence production frequencies are presented in Table 3

(Insert Table 3 about here)

A logistic regression using base processing instruction and solution principle as predictors yielded a significant model, $\chi 2(1, N = 72) = 5.9$, $p = .05$, with base processing being the only significant predictor in the model, Wald = 5.51, $p = .019$. Deriving a solution principle did not affect production of the dispersion-convergence solution.

There were no differences between the conditions in the number of words produced in the summaries of the critical base information. However, participants who evaluated the base information produced significantly fewer solution components in their summaries ($M = 2.14$, $SD = 0.59$) than participants who had not evaluated the base information ($M = 2.44$, $SD = 0.56$), $t(70) = 2.32$, $p = .03$.

In contrast to Mandler and Orlich (1993), 50% of participants in this study were able to follow the summarizing instructions successfully. Nonetheless, only 25% of participants produced the correct dispersion-convergence principle in their summaries, and production of the correct solution principle did not predict task success on the radiation problem (4 out of 9 solvers) compared to other solution principles (again 4 out of 9 solvers).

Discussion

The detrimental effect of evaluative processing was replicated again in Experiment 1c. Asking participants to produce a solution principle to the base information did not improve transfer. It also appears that time spent encoding the base information seems not to affect rates of transfer since participants spent nearly ten minutes longer on the encoding stage in Experiment 1c than in Experiment 1a and 1b with no additional benefit to analogical problem solving.

Analysis of the summaries showed that participants who evaluated the base information produced fewer solution components, an effect identified in the post-transfer base-story recall data of Experiment 1b. These data further suggest that evaluation causes an impoverished representation of the base information, and this impoverished representation is resistant to attempts to encourage participants to extract the appropriate information by summarization.

In Mandler and Orlich's study a third of participants who summarised correctly (15/45) produced the dispersion-convergence solution pre-hint. In the present study only 11% of participants produced the dispersion-convergence solution pre-hint, only a very small increase (3%) on Experiment 1a where participants were not instructed to produce a summary. This experiment is somewhat different to Mandler and Orlich's in that the summary instructions were given to participants after they had processed the base material once, rather than on initial presentation. Also, in the present study, a less obviously analogous solution than that used by Mandler and Orlich was placed amongst two distracter solutions. These manipulations may act to reduce the benefits of producing a summary of base information.

Experiment 2

Positive effects of *explanation* have typically been found with base problems that are both superficially and conceptually similar to the target problem (Chi et al., 1989; Chi et al., 1994). The nuclear power station story employed in the experiments so far has conceptual but not superficial similarity to the target radiation problem, so

Experiment 2 employed a base story used previously in the analogy literature, namely Holyoak and Koh's (1987) "Light Bulb" story, which is both superficially and conceptually similar. The Light Bulb story is structurally similar to the target radiation problem in that full power ultrasound cannot be employed to repair a broken filament in a light bulb because it will break the fragile glass surrounding the filament. The solution is to have many ultrasound machines set at low intensity surrounding the light bulb that converge to create a high intensity dose only at the point of the filament. This solution has superficial similarity to the radiation problem in that a ray-producing machine is being used to perform a delicate operation. The Light Bulb story is thus superficially as well as conceptually similar to the target radiation problem. The experiment also provided an opportunity to see if the negative effects of evaluation can be overcome by requiring participants to generate explanations of the base problem and solution.

Method

Design. There were two between-participants factors, Evaluation (present or absent) and Explanation (present or absent). Dependent variables were as in earlier experiments.

Materials. The materials used in Experiment 1a were used here with the exception that the Light Bulb story (Holyoak & Koh, 1987, fragile glass and ultrasound version) was used as the analogous base. Holyoak and Koh found that using the Light Bulb story as a base yields pre-hint target problem solving performance of 38%.

Participants. 63 female and 33 male students from Lancaster University participated in the experiment. Participants were paid £4 for the experiment, which took approximately 40 minutes. Participants were randomly allocated to the different conditions.

Procedure. The same procedure as Experiment 1a was used (without the memorization instructions) except that the factor of explanation was added. In the explanation conditions, participants attempted to explain the story and the solution as if they were talking to an alien. The participants in this condition were informed that, "The alien understood English but had no common sense whatsoever". This is similar to instructions used by Galotti and Gangon (1985). Participants in the evaluation/explanation condition were instructed at the beginning of the task that they were to perform both processes but were told neither how much time to allocate to each process nor which order they should perform them in. If participants lapsed into silence for more than five seconds they were asked to keep talking. After attempting the target problem, participants were asked to recall the light bulb base story.

Results

The frequency of dispersion-convergence solution production pre- and post-hint is shown in Table 4.

(Insert Table 4 about here)

A logistic regression of the pre-hint data using evaluation instruction and explanation instruction as predictors yielded a significant model, $\chi 2(1, N = 96) = 7.7$, $p = .005$, with evaluation instruction being the only significant predictor, Wald = 7.33, $p = .007$. A logistic regression of the post-hint data using evaluation instruction and explanation instruction as predictors yielded a significant model, $\chi 2(1, N = 96) = 14.4$ (2), $p = .001$, with the only significant predictor once again being evaluation instruction, Wald = 11.38, $p = .001$. Significantly fewer elements of the solution were produced by participants in the evaluation conditions ($M = 2.54$, $SD = 1.87$) than non-evaluation groups (*M* = 3.57, *SD* = 1.36), $F(1, 96) = 7.92$, $p = .006$. Neither explanation nor the interaction between evaluation and explanation were significant. Participants in the evaluation conditions also recalled fewer solution elements of the base story (*M* = 1.94, *SD* = 0.75,) than participants in non-evaluation conditions (*M* = 2.26, *SD* = 0.83), $F(1, 96) = 4.20$, $p = .04$.

Discussion

The detrimental effect on transfer caused by an evaluation instruction was reproduced in Experiment 2 with different materials and a different relation to that used previously (i.e., conceptually and superficially similar, rather than conceptually similar and superficially distinct). Moreover, the negative effect of evaluation was observed not only post-hint but also pre-hint. Somewhat surprisingly, explanation did not facilitate transfer. One reason may be that the two minutes processing time in this experiment was inadequate for participants to derive the benefits of explanation, particularly when the time was shared between evaluation and explanation.

Thus it appears that evaluation has a consistently detrimental effect on analogical transfer. Experiments 3 & 4 were designed to explore further why this phenomenon occurs.

Experiment 3

Experiment 3 enabled an examination of the verbal protocols produced during participants' processing of the base material. This experiment also amended the procedure so that the solution was presented to participants along with the story, thus removing the apparently random selection process used in Experiments 1 & 2. This change was made to address an argument that participants in the evaluation conditions of the previous experiments might have devalued the dispersion/convergence solution because it was selected at random.

Method

Design. A one-factor experiment was run with base processing (evaluative, explanatory, summary) as the independent variable. Dependent variables were as in Experiment 1b, with the addition of two measures of extraneous base information processing (sentence segments and percentage of time).

Materials. The materials were the same as those used in Experiment 1 with the exception that the solution to each of the base stories was included as part of the story.

Participants. 21 female and 5 male students from University of South Australia participated in the study. Participants received a $20 book voucher for participating in the experiment, which took approximately 30 minutes. Participants were randomly allocated to the different conditions.

Procedure. The procedure of Experiment 1a was employed without the memory instruction and with the addition of an evaluation condition. Participants were allocated to one of three conditions: evaluation, explanation or summary. Instructions for these conditions were as in previous experiments.

Results and Discussion

No participants produced the dispersion-convergence solution to the radiation problem pre-hint and three participants produced the dispersion-convergence solution post-hint (two in the explanation condition and one in the summary condition). The number of elements of the target solution (x-rays, convergence, multiple, lower intensity) produced by participants in the three conditions was marginally significant, $F(2,23) = 2.56$, $p = .096$, with mean differences in the expected direction (Evaluation $M = 1.67$, $SD = .866$; Explanation $M = 2.44$, $SD = 1.13$; Summary $M = 2.63$, $SD = .74$). These mean differences are similar to those observed in previous experiments. Recall of the components of the base material was non-significant, although again mean differences were in the expected direction (Evaluation *M* = 1.67, *SD* = 1.23; Explanation *M* = 2.44, *SD* = .527; Summary *M* = 2.25, *SD* = .463) and were similar to mean differences observed in previous experiments. While the mean differences were not statistically significant (probably due to the smaller number of participants used in this experiment because of the full verbal protocol analysis) this experiment does appear to be comparable to the other experiments presented in this paper.

A qualitative analysis of verbal protocols produced by participants in the evaluation condition suggest that participants paid continuous attention to the problem and understood how the solution solved the problem. For example, one participant stated "they can't use just one pipe because of the pressure so to have several pipes is a good idea." However, participants in the evaluation condition also often considered issues that were irrelevant to the structure of either the problem or solution. For example, one participant considered whether the solution would take water away from other sources, while another considered the potential lack of building resources (a full transcript of the verbal protocols of these participants is presented in Figure 1). This suggests that one of the reasons for the detrimental effect of evaluation is that participants are considering extraneous information while evaluating the base stories, which is interfering with their recall of the critical solution information.

To capture the amount of extraneous information considered by participants, two measures were calculated. These measures were the number of sentence segments in the verbal protocols that were not in the analogous base story and the percentage of verbalisations pertaining to information that was not related to the story or solution. Participants in the evaluation condition were significantly more likely to discuss extraneous information during base processing (sentence segments *M* = 7.33, *SD* = 2.5; time *M* = 65.6, *SD* = 17.5) compared to either explanation (sentence segments $M = 2.89$, $SD = 2.8$; time $M = 21.8$, $SD = 25.3$) or summary (sentence segments $M = 2.63$, $SD = 1.9$; time $M = 15.9$, $SD = 18.1$), $F(2,23) = 10.4$, $p = .001$ for sentence segments and $F(2, 23) = 14.98$, $p < .0001$) for time. The amount of extraneous information in participants' verbal protocols (as measured by both sentence segments and time) was also significantly negatively correlated with recall of the elements of the dispersion-convergence solution in the story recall task (sentence segments $r = -.52$, $n = 26$, $p = .007$; time $r = -.47$, $n = 26$, $p = .015$). This suggests that the increased amount of extraneous information verbalised in the evaluation condition is leading to a representation of the base information that makes the critical solution elements more difficult to recall. The correlations between the measures of extraneous base information and number of correct elements of the solution were in the predicted direction but were not significant (sentence segments $r = -.14$, $n = 26$, $p = .5$; time $r = -.13$, $n = 26$, $p = .53$)

A potential explanation for the negative effect of evaluation found in these experiments is that participants in the evaluation conditions are rejecting the dispersion-convergence solution during base processing which might then cause them to not consider dispersion-convergence solution as a viable solution to the radiation problem. To explore this explanation, participants' evaluations of the effectiveness of the dispersion-convergence solution in the *evaluation condition* were collected from their verbal protocols for the analogous base story. Six participants thought the dispersion-convergence solution was effective (three participants used the words good or best solution, one participant stated that the solution was what they would do, one participant stated that the solution was fairly achievable and one said it would allow it [the Nuclear power station] to cool down). Another participant stated that it would be better to have a number of pipes but it was unclear whether they were simply paraphrasing the story. Only two participants produced a negative evaluation of this solution (one participant said it was not reasonable and another stated that it [the dispersion-convergence solution] wouldn't make much difference). One other participant thought there were better solutions that didn't involve nuclear power, but did not produce a negative evaluation of the dispersion-convergence solution. The other participant did not evaluate the dispersion-convergence solution instead focusing on why a single pipe would rupture.

The ideas that evaluation leads to solution rejection is also unsatisfactory given that the time allocated for solving the radiation problem in all of the experiments was usually sufficient to exhaust participants' solutions, and participants frequently produced solutions that they didn't think would work (such as operating on the patient, using positive thoughts or freezing the patient so they could be operated on in the future). In addition, in Experiment 1b, participants were encouraged to produce as many solutions as they could think of regardless of whether they thought they were good solutions to the problem. It would be expected that if participants were simply rejecting the dispersion-convergence solution as an ineffective solution they would produce it as a solution under these conditions. This suggests that the negative effect of evaluation is not due to participants rejecting the dispersion-convergence solution during base processing which might then cause them to not consider dispersion-convergence solution as a viable solution to the radiation problem.

Experiment 4

The finding that participants in the evaluation condition are producing significantly more extraneous information during their base processing and that this is negatively correlated with story recall gives rise to two potential explanations. One explanation is that because evaluation encourages participants to explore a wider range of material than other manipulations it leads to a failure to attend adequately to the to-be-transferred information. The other is that considering a wide range of information is leading to a mental representation where extraneous information interferes with access to the to-be-transferred information.

Thus, Experiment 4 was designed to mediate between two possible explanations for the negative effect of evaluation: inattention and interference, by requiring participants either to evaluate the base information first and then explain it or to explain the base information first and then evaluate it. If the negative effect of evaluation arises through inattention, then an explanation task should complete encoding of the base problem regardless of the position of evaluation. Alternatively, the structure of the problem representation may be fixed by an initial evaluation task in such a way as to limit or otherwise interfere with participants' access to relevant problem information. Thus, if the negative effect of evaluation arises through interference, then evaluating before explaining should lead to lower levels of analogical transfer than the opposite task order.

Method

Design. A one-factor experiment was run with position of evaluation (first or second) as the independent variable. Dependent variables were the same as those used in Experiment 2.

Materials. The materials were the same as those used in Experiment 2.

Participants. 23 female and 13 male students from Lancaster University participated in the experiment. Participants were paid £4 for the experiment which took approximately 45 minutes. Participants were randomly allocated to the different conditions.

Procedure. The basic procedure of Experiment 1a was employed. Participants were informed that they would be required to undertake two processes: evaluation and explanation. In one condition participants were instructed to evaluate the problem and solution and then explain it. In the second condition participants were instructed to explain the problem and solution and then evaluate it. Participants spent approximately one minute evaluating each base story with the rest of the time spent explaining it. Participants spent three minutes on each base story rather than the two minutes employed in previous experiments[1]. If participants lapsed into silence they were encouraged to keep talking.

Results

The frequency of dispersion-convergence solution production pre- and post-hint is shown in Table 5.

(Insert Table 5 about here)

Both pre-hint and post-hint dispersion-convergence production was significantly different between groups: pre-hint, $\chi 2(1, N = 36) = 5.2$, $p = .02$; Wald = 4.33, $p = .03$; post-hint $\chi 2(1, N = 36) = 9.7$, $p = .002$; Wald = 8.14, $p = .004$. The number of elements of the target solution (x-rays, convergence, multiple, lower intensity) produced by participants in the evaluation-then-explanation condition (M = 1.82, SD = 1.86) was significantly lower than that produced by participants in the explanation-then-evaluation condition (M = 3.72, SD = 0.97), $t(34) = -3.80$, $p = .001$). Participants in the evaluation-then-explanation condition also recalled significantly fewer solution elements (convergence, multiple and lower intensity) from the base problem (M = 1.52, SD = 1.20) than participants in the explanation-then-evaluation condition (M = 2.33, SD = 1.11), $t(32) = -2.10$, $p = .05$.

Discussion

The results support an interference explanation of the negative effects of evaluation. If solvers form a representation of the base problem through explanation, then this appears to protect them against effects of the evaluation process. However, if participants evaluate before they explain then they form a representation from which the to-be-transferred information is obscured. Differences in recall of the base problem between the two conditions indicate that the effect is not simply inattention to relevant information: evaluating first appears to set the structure of the representation in such a way as to limit the subsequent effectiveness of explanation.

General Discussion

In four experiments, a task requirement to evaluate the solution to a base problem was shown to have a consistently detrimental effect on transfer to an analogous target problem. This effect generalised across different base stories and was not ameliorated by asking participants to summarise the base information for a later memory test (Experiment 1a), to engage in evaluation tasks with both base and target problems (Experiment 1b), or to elaborate the dilemma, goal and solution principle of the base problem (Experiment 1c). When participants explained the base problem and solution at the same time as evaluating it, explanation did not have the positive effect that is generally reported in transfer experiments (Experiment 2). However, when explanation preceded evaluation, it removed the negative effect of evaluation on analogical transfer (Experiment 4). The outcome of these experiments was completely unexpected. Our initial hypothesis was that evaluation would act to elaborate participants' mental representations of base material, which in turn should enhance retrieval and mapping of the base problem and solution during solution of the target problem. One possible explanation of the results is that participants evaluated the solution to the base problem as unsatisfactory, and therefore rejected it as a solution to the target problem. However, participants' comments on the solution in Experiment 3 were largely positive with only 2 participants considering the dispersion-convergence solution to be unsatisfactory.

Another possibility is that, in contrast to initial expectations, evaluation did not lead participants to elaborate their understanding of the base problem and its solution. Again, inspection of participants' verbalisations suggests otherwise. As Figure 1 exemplifies, participants processed the base problem and solution in a fine-grained way, examining all elements: the pipes, the materials needed for the pipes, the pumping station, etc. Moreover, their evaluations were not facile. Using Chi's (2000) description of an inference and analysing the transcripts at the level of the sentence segment[2] (which corresponds to the idea segments of Chi et al., 1989), we counted the number of inferences generated by participants in Experiment 1a. Thirty three participants who evaluated the base information produced 279 segments of talk, 163 of which were inferences, the remainder being paraphrases of base information, self-monitoring statements or restatements of inferences. This average of 58% inferences in the protocols compares well with the 29% inference generation that Chi et al. (1989) considered to be indicative of high self-explainers.

The finding concerning evaluation in this study is reminiscent of the effects of asking participants to justify their choices (Sieck, Quinn, & Schooler, 1999). Sieck et al. found that participants who provided a written justification of their prospective story-matches focused on superficial rather than structural aspects of the matches compared to participants who solved a cross-word puzzle. Sieck et al. argued that this occurs because structural information is harder to verbalize than superficial information so people will essentially ignore structural information and focus instead on superficial commonalities. Drawing a parallel between this work and the present study would suggest that participants in the evaluation condition are focusing on superficial elements at the expense of structural elements because structural elements are hard to verbalise. This explanation is unsatisfactory for a number of reasons. First, there is no reason to suggest a priori that the structural elements of the base information are harder to verbalise in the evaluation condition than in either the summarizing or explanation conditions. Second, it is clear that the effect of evaluation is to allow participants to consider extraneous information relative to other conditions. Third, a verbal overshadowing explanation provides a less parsimonious explanation for these data than one which assumes that evaluation impairs performance because *too much* information is activated during processing. Furthermore, it should be noted that the concept of justification used by Sieck et al. has a rather different meaning to the concept of evaluation used here, as can be discerned from the fact that Sieck et al. also refer to "explanation" as justification. As we have found in this series of experiments, explanation and evaluation have very different effects on performance.

Instead, it appears that participants in the evaluation condition *over-elaborated* the base problem and solution. Participants in the evaluation condition produced significantly greater amounts of extraneous information than participants in other conditions during base processing, which was associated with impaired recall of the critical solution elements (Experiment 3). The task of *explanation* appears to focus participants on the specific interconnected body of knowledge presented in the base problem, while *evaluation* appears to encourage the participants to consider wider issues outside the information presented. Therefore, inferences that are made when participants evaluate lead to the development of a mental representation of the base information that contains extraneous information that is not relevant for solving the later target problem[3].

Over-elaboration might lead participants to fail to attend to relevant features in constructing a mental representation of the base problem and solution. Alternatively, it may act to fix the mental representation in such way that participants cannot access relevant information once the initial representation is set. The results of Experiment 4 provide provisional support to the latter hypothesis: participants who explained the base problem and its solution before they evaluated did not show the negative effect of evaluation, whereas participants who evaluated first and explained second showed the effect. If the effect of evaluation was simply to change the elements that participants attend to, then one might expect that it would have a negative effect regardless of its task position. Instead, it appears that the critical limitation arises when it comes first, indicating that it is the impact on initial formation of mental representation that is critical. Formation of a flexible initial representation of the source problem appears to be a critical pre-condition for successful analogical problem-solving, a process which can be interfered with by evaluation.

The results of these experiments may also have practical implications. Neuman and Schwartz (1998) warn that “when considering prompting people to self-explain, in order to foster learning/problem solving, the content of the self-explanations should be carefully examined.” (p. 22). We suggest that potential problems with premature evaluation should be considered when using explanation as a pedagogic method. When an instructor asks “What do you think of this solution?” they may be unwittingly impacting negatively on a student’s understanding of a problem, when a slightly different question “What do you think this solution is about?” might yield a more beneficial outcome. The current experiments employed a forced evaluation procedure, and it is not clear whether spontaneous evaluation will also impair analogical transfer. Nonetheless, it seems possible that the imposition of a task requirement to evaluate can limit the ideas people generate during creative problem solving and decision making. Creative thinking techniques such as brainstorming that advocate the postponement of evaluation while ideas are being created may allow individuals to optimise their initial representations of problem domains. Considering that an ability to evaluate the qualities of solutions is generally considered to be one of the cornerstones of intelligent thought, a negative effect of evaluation has surprising implications for the methods used to encourage successful problem-solving and creative thought.

References

Anoli, L., Antonietti, A., Crisafulli, L., & Cantoia, M. (2001). Accessing source information in analogical problem-solving. *Quarterly Journal of Experimental Psychology, 54A*, 237-261.

Ashton, R. H., & Kennedy, J. (2002). Eliminating recency with self-review: The case of auditors 'going concern' judgements. *Journal of Behavioural Decision Making*, 15, 221-231.

Bearman, C., Ormerod T. C., & Ball, L. J. (2003). A negative effect of evaluation on analogical problem solving. In R. Alterman, & D. Kirsh (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (pp. 121-126). Mahwah, NJ: Lawrence Erlbaum Associates.

Blessing, S. B., & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 792-810.

Catrambone, R., & Holyoak, K.J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1147-1156.

Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M.T.H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Darses, F. (2002). A framework for continuous design of production systems and its application in collective redesign of production line equipment. *Human Factors and Ergonomics in Manufacturing*, 12, 55-74.

Dunbar, K. (2001). The analogy paradox: Why analogy is so easy in naturalistic settings, yet so difficult in the psychological laboratory. In D. Gentner, K. J. Holyoak, & B. Kokinov (Eds.), *Analogy: Perspectives from cognitive science* (pp. 313-334). Cambridge, MA: MIT Press.

Dunker, K. (1945). *On problem solving*. Westport: Greenwood Press.

Easton, G. (1992). *Learning from case studies (2nd Edn.)*. London: Prentice Hall.

Galotti, K. M., & Gangon, III, W.F. (1985). What non-programmers know about programming: Natural knowledge procedure specification. *International Journal of Man-Machine Studies, 22*, 1-10.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Psychology, 7,* 155-170.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393-408.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15,* 332-340.

Keane, M. T. (1996). On adaptation in analogy: Tests of pragmatic importance and adaptability in analogical problem solving. *Quarterly Journal of Experimental Psychology, 49A*, 1062-1085.

Klein, G. (1998). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.

Mandler, J. M., & Orlich, F. (1993). Analogical transfer: The roles of schema abstraction and awareness. *Bulletin of the Psychonomic Society, 31*, 485-487.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structured of behaviour*. New York: Holt, Reinhhart & Winston.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer, memory-oriented training promoted memory for training. *Memory and Cognition, 19*, 543-557.

Neuman, Y., & Schwartz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem solving. *British Journal of Educational Psychology, 68*, 15-24.

Novick, L. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 510-520.

Ormerod, T. C., Chronicle, E. P., & MacGregor, J. N. (2006). Asymmetrical analogical transfer in insight problem solving. In R. Sun & N. Miyake (Eds.), *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society* (pp. 1899-1905). Alpha, NJ: Sheridan Printing.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. *Memory and Cognition, 29,* 214-221.

Pirolli, P., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology, 39*, 240-272.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043-1056.

Sieck, W. R., Quinn, C. N., & Schooler, J. W. (1999). Justification effects on the judgment of analogy. *Memory and Cognition, 27,* 844-855.

Schunn, C. D., & Dunbar, K. (1996). Priming, analogy and awareness in complex reasoning. *Memory and Cognition, 24*, 271-284.

Van Lehn, K., & Jones, R. M. (1993). What mediates the self-explanation effect? Knowledge gaps, schemas or analogies? In M. Polson (Ed.), *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1034-1039). Mahwah, NJ: Lawrence Erlbaum & Associates.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129-140.

Author Note

The first author's research was funded by an ESRC CASE award to Prof. Ormerod, with co-funding from the European Case Clearing House. Experiments 1a and 1b formed the basis of the following presentation: Bearman, C. R., Ormerod, T. C. and Ball, L. J. (2003), A negative effect of evaluation upon analogical problem solving. *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: LEA.

.

Footnotes

[1]A pilot study found that increasing the base evaluation time from two minutes to three minutes had no effect on production of the dispersion-convergence solution (2 mins = 2/9 pre-hint and 5/9 post-hint; 3 mins = 2/9 pre-hint and 7/9 post-hint; $\chi 2$ = 0.1, ns, Wald = 0.10, ns).

[2]Sentence segments are segments of talk based around sentences or sections of sentences parsed at the combination words such as ‘and’, ‘but’ or ‘erm’

[3]The participants do not, of course, know what will be relevant for the later task or even that there will be a task that is dependent on the base information.

Appendices

The Nuclear Power Station Story

A water pumping station in the hills above Peshawar has recently been adapted to feed a nuclear power facility, where the water acts as a coolant. Unfortunately the main pipe that leads from the water pumping station to the nuclear facility cannot stand the increased pressure and will eventually rupture. The nuclear power facility must have a continuous supply at a certain volume or it will overheat and go into meltdown. Experts have found that it is impossible to build a single large pipe to carry the water with the materials available that will not also rupture due to the high pressure of the water flowing through it. The suggestion is: to build a number of pipes that have less water pressure.

The Osteopath Story

After many years of research, a Philadelphia doctor of Osteopathy thinks that he has found a way to alleviate chronic back pain. The doctor theorizes that if a high velocity thrust can be applied to a fairly wide section of the lumber region of a person's back they will be cured. Unfortunately, the doctor has found it impossible to test his theory because he is unable to deliver a high velocity thrust of sufficient force because such a thrust requires more strength than the doctor possesses. The suggestion is: to have the doctor develop a machine that is capable of delivering the high velocity thrust

The Commander Story

A military government was established after the elected government was toppled in a coop. The military imposed martial law and abolished all civil liberties. A tank corps commander and his forces remained loyal to the overthrown civilian government. They hid in a forest waiting to launch a counter attack. The commander felt that he could succeed if only the military headquarters could be captured. The headquarters was located on a heavily guarded island situated in the center of a lake. The only way to reach the island was by way of several pontoon bridges that connected it to the surrounding area. However, each bridge was so narrow and unstable that only a few tanks could cross at once. Such a small force would easily be repulsed by the defending troops. The headquarters therefore appeared to be invincible. The suggestion is: to get the commander to have his forces build a bigger bridge that will allow a sufficient force to cross to the island.

The Light Bulb Story

In a physics lab at a major university, a very expensive light bulb that would emit precisely controlled quantities of light was being used for some experiments. Ruth was the research assistant responsible for operating the sensitive light bulb. One morning she came into the lab and found to her dismay that the bulb no longer worked. She realized that she had forgotten to turn it off the previous night. As a result the light bulb overheated and two wires in the filament inside the bulb fused together. The surrounding glass bulb was completely sealed, so there was no way to open it. Ruth knew that the light bulb could be repaired if a brief, high-intensity ultrasound wave could be used to jar apart the two fused parts. Furthermore, the lab had the necessary equipment to do the job. However, a high-intensity ultrasound wave would also break the fragile glass surrounding the filament. At lower intensities the ultrasound wave would not break the glass, but neither would it jar apart the fused parts. So it seemed that the light bulb could not be repaired and a costly replacement would be required. The suggestion is: to place several ultrasound machines in a circle around the light bulb and administer low-intensity ultrasound waves from several directions at once.

The Radiation Problem

A doctor is faced with a patient who has a malignant tumour in his stomach. It is impossible to operate on the patient, but unless the tumour is destroyed the patient will die. X-rays can be used to destroy the tumour. If an X-ray reaches it at sufficiently high intensity the tumour will be destroyed. Unfortunately, at this intensity the healthy tissue that the ray passes through on the way to the tumour will also be destroyed. At lower intensities the ray is harmless to healthy tissue, but it will not affect the tumour either. It looked like the patient was going to die.

Participant 30 - Erm yeah I think this is a good idea. Erm because obviously they can't use just one pipe because of the pressure so to have several pipes is a good idea because I don't think there's any other possible way and the nuclear facility has to have this increased pressure all the time so obviously all these pipes are going to do the job because they will have the required amount of water going through you know keeping it cool what have you erm yeah I suppose the problem is whether the pumping station at the top will be able to you know cope with the amount of water it is pumping out but I suppose they will take all that into consideration when they get their little engineers out [laughter] and sort of planning it all and everything so yeah so its an ok suggestion to me [laughter].

Researcher - OK, several more seconds.

Participant 30 - Oh err yeah as long as its not going to sort of you know draw water from you know obviously its going somewhere else as well as long as its not taking it away from another source then its you know it should be fine [laughter].

Researcher - Ok, that's fine

Participant 11 - Building several pipes might be able to handle the increased water because you could have as many pipes as you want, but, you might still be limited by, the fact that they've only got limited materials available. Because it said that, with the that it's impossible to build a single large pipe with the materials available. Umm so it would depend on whether the materials available were suitable for building. And on the small points and other than that it sounds like quite a good solution.

Researcher - You've got a few more seconds if you would like to continue.

Participant 11 - [Laugher] suppose you'd have to work out how many pipes you needed, to handle, so that each one would have less water pressure. You didn't have enough you could still get the pipe fracture?

Figure 1

Two Examples of Participants Evaluating Base Information

Table 1

Frequency of the Production of Dispersion-Convergence Solutions Post-Hint by Condition in Experiment 1a

	Base processing instruction	
Memory instruction	Evaluative encoding	Non-evaluative encoding
Memory encoding	5/18 (28%)	11/18 (61%)
No memory encoding	4/18 (22%)	9/18 (50%)

Table 2

Frequency of the Production of the Dispersion-Convergence Solution Post-Hint by Condition in Experiment 1b

	Base processing instruction	
Target processing instruction	Evaluative encoding	Non-evaluative encoding
Evaluative processing	4/18 (22%)	9/18 (50%)
Non-evaluative processing	3/18 (16%)	9/18 (50%)

Table 3

Frequency of the Production of the Dispersion-Convergence Solution Post-Hint by Condition in Experiment 1c

	Base processing instruction	
Solution principle	Evaluative encoding	Non-evaluative encoding
Generated	4/18 (22%)	10/18 (55%)
Absent	6/18 (33%)	11/18 (61%)

Table 4

Frequency of Production of Dispersion-Convergence Solutions Pre-Hint and Post-Hint by Condition in Experiment 2 (each cell out of 24)

	Pre-hint		Post-hint	
	Evaluation instruction		Evaluation instruction	
Explanation instruction	Present	Absent	Present	Absent
Present	6 (25%)	12 (50%)	12 (50%)	17 (71%)
Absent	5 (21%)	12 (50%)	11 (46%)	23 (96%)

Table 5

Frequency of Production of Dispersion-Convergence Solutions Pre-Hint and Post-Hint by Condition in Experiment 3 (N = 18)

Position of evaluation instruction	Pre-hint	Post-hint
Explanation-then-evaluation	8 (44%)	15 (83%)
Evaluation-then-explanation	3 (17%)	6 (33%)