Title

Using instruction checks to measure source understanding in analogical transfer of insight solutions

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Ross, Wendy
Ormerod, Thomas C

Publication Date

2024

Using instruction checks to measure source understanding in analogical transfer of insight solutions.

Wendy Ross (w.ross@londonmet.ac.uk)
Department of Psychology, London Metropolitan University,
London, N7 8DB UK

Thomas C Ormerod (t.ormerod@sussex.ac.uk)
School of Psychology, University of Sussex, Pevensey 2 5b5,Falmer,
East Sussex, BN19QH, UK

USA

Abstract

Analogical transfer between source and target problems ought to be a major contributor to problem-solving and learning. Yet, data from laboratory studies show that successful spontaneous analogical transfer does not reliably occur in the absence of explicit hints to analogize, in the presence of a delay between source and target, or when there are extensive filler tasks, a finding attributed to the complexity of analogy retrieval and mapping. Here, we show that participants solving variants of the Cards problem often failed to show transfer between source and target problems that shared both conceptual and superficial similarities. Frequency of re-inspecting the task instructions was a significant predictor of transfer, with participants successful at T2 requiring fewer re-inspections. The results suggest that analogical transfer may be limited, not just by the difficulty of mapping between source and target, but by a lack of conceptual understanding of the source and its solution, even when the source is solved.

Keywords: Analogical transfer; insight; problem-solving; instructions.

Problem-solving has been a key part of cognitive research since the foundations of the field of cognitive science in the 1950s. The parameters of the problem-solving experiment are quite simple. Participants are given a problem – often knowledge-lean (i.e., with the problem statement containing everything needed to solve the problem) and well-structured (i.e., where a space of possible states between the initial and goal states can be specified) – and they are tasked with generating a correct solution to the problem. Performance is typically measured in terms of success at solving the problem and latency to solution. Experimental research consists of manipulating a set of theoretically justified factors to see the extent to which performance is sensitive to these changes (Gozli, 2017). In the case of analogical transfer, it is generally assumed that success in solving one or more source problems provides the basis for analogical transfer to a conceptually related target problem (see Gick & Holyoak, 1980 among others).

In analogical transfer, a problem is solved when its solution draws on understanding gained from solving or being presented with solutions to conceptually similar problems. The classic example of this is transfer between Duncker's (1945) Radiation problem[1] and its analogical variants (e.g., the General and Red Adair analogs – Gick & Holyoak, 1980). In a series of experiments by Gick & Holyoak (1980; 1983), they demonstrated a set of conditions under which adult participants were able to transfer solution knowledge between superficially different but conceptually similar variants of the Radiation problem. Notably, analogical transfer is elicited when participants are given a hint as to the analogical structure of source and target problems, when they receive multiple source problems plus solutions prior to attempting a target problem, and when they receive a schematic or abstract description of the underlying problem and solution structure prior to attempting the target. In other words, transfer becomes possible when conceptual understanding of the source and target problem is established, and this process often requires external intervention.

The early success of analogical transfer studies led some cognitive theorists (e.g., Anderson, 2014) and educationalists (e.g., Kolodner, 1997) to position analogical reasoning as the fundamental basis of many human reasoning and learning capacities. Despite these successes, analogical transfer remains remarkably hard to elicit. Indeed, as in the case of the studies of Gick and Holyoak, there are few demonstrations of *spontaneous* analogical transfer: instead, successful transfer, in laboratory studies at any rate, seems to emerge only when participants are given a strong hint to analogize between problems, or when source

[1]Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient but unless the tumor is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once with sufficiently high intensity the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but they will not affect the tumor either. How might the tumor be removed without harming the patient further?

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

and target problems are presented with little or no intervening time or alternative stimuli.

Catrambone and Holyoak (1989) demonstrated that it is possible to generate some degree of spontaneous transfer with a gap of one week between source and target problem-solving, if participants undertake a comparison of two source analogs (Expts.2 and 3) and the wording of the target problem is adapted to be more similar to the wording of the source problems (Expt. 3), though the effects are modest (for comparison groups with a one-week delay, Expt. 2 yielded 7% solvers from n = 15 while Expt. 3 yielded 23% solvers from n = 13, compared with an expected base rate of 10% solvers). Higher levels of target solutions were found in Experiment 5 (64% solvers from n = 11) when participants were given a rigorous comparison script that they applied to three source analogs. The intensity of the manipulation required to elicit these levels of transfer calls into question whether it can be called 'spontaneous', but also shows the importance of source comprehension to transfer.

Explanations for the lack of transfer tend to focus on two factors. First, unless source and target problems share superficial similarities, such that the thematic content or representation of the target can act as retrieval cues, the relevance of an analogical source may not be recognized (e.g., Reeves & Weisberg, 1994). Second, the task of mapping the analogical relations between source and target (i.e., identifying what is a conceptual structure within the source and how it applies to the target) can be complex (e.g., Keane, Ledgeway, and Duff, 1994). Ormerod (2023) has argued that the lack of default use of analogy should be no surprise: recognizing that two problems are analogous can be cognitively challenging and error prone, and, given the possibility of solving a problem without recourse to analogy, may be deemed by the solver to be unnecessary or inefficient. Where the analogical nature of source and target is recognized, and where mapping of relations can take place, there is an assumption that conceptual knowledge of the problem has been gained during solution of the source: If you solve it, then surely you understand it.

One possibility that has not, to our knowledge, been examined is that analogical transfer may not arise because, even if participants have solved or are presented with the solution to the source problem, their understanding of the source problem itself may be insufficient to enable them to map its solution onto the target problem. In essence, they may solve the source problem without insight into the nature and route to production of the solution itself. This is consistent with the findings of Catrambone and Holyoak (1989) which, in summary, seem to suggest that the more work that is put into understanding the nature of the source problem(s), the more likely spontaneous transfer is to occur.

To date, research in problem-solving has been driven by the unexamined assumption that successful problem-solving is uniform in nature and underpinned by the same level of conceptual understanding. There are demonstrations that enhancing the understanding of a source problem beyond mere solution can enhance transfer. For example, Ormerod and MacGregor (2017) explored solutions to the Nine-ball problem, a task in which participants must discover which of nine otherwise identical balls is fractionally heavier, using a balance scale only twice. They found that adding a solution-irrelevant constraint that each ball cost £1 to weigh and that only £8 was available in total to solve not only increased solution rates to the source problem but also increased rates of spontaneous transfer to problem analogs relative to participants in a condition who had £12 to solve. They suggest that the additional constraint focused participants' understanding on key aspects of the source problem. The current experiment investigates whether similar changes in source understanding can arise without adding further constraints to the source problem.

The Cards Problem

In the study that follows, we chose not to use the Radiation problem, partly due to its familiarity among participant pools, but also because radiation technology for cancer treatment is more widely understood, changing the base rate of problem difficulty. Instead, we chose to focus on the less well-known 'Cards' problem (Cunningham & MacGregor, 2008) in which the task is to lay in a grid the Royal Cards (or other subsets of cards) from a pack of playing cards in such a way that there is exactly one card of each denomination in each row and column of the grid. Solution examples are shown in Figure 1.

To meet these constraints, participants must build a grid that includes empty spaces, as in Figure 1(a). The problem statement does not indicate any need for spaces between cards, and the solver must discover this concept. Thus, the Cards problem may be conceived as one likely to yield the phenomena of insight problem-solving. It is analogous to the puzzle faced by Mendeleev in his synthesis of the periodic table of chemical elements. Mendeleev's breakthrough included the insight that elements could be arranged according to their known properties in a grid of rows and columns only if spaces were inserted for as-yet-to-be discovered elements (Akin & Akin, 1998).

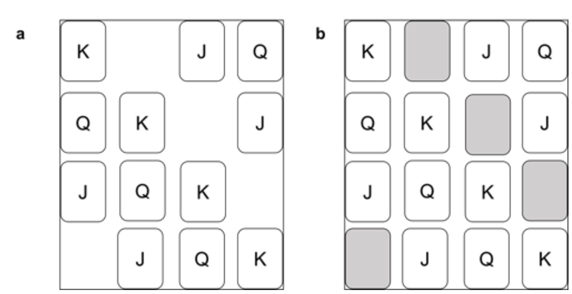

Figure 1: Solution to the card problem (a) without blank card hints and (b) with blank card hints.

If participants are instead given 16 cards (4 Jacks, 4 Queens, 4 Kings and 4 Blanks), and told that the task remains the same (to place the Royal Cards uniquely in each row and column using the blanks as guides), the problem changes. The blank cards function as hints to leave spaces between the Royal cards (Figure 1b), allowing the problem-solver to lay out a problem solution without needing to discover the 'insight' to leave gaps between cards. The addition of the blank cards yields hint and non-hint versions of the problem. Given that the blank cards may effectively act as gaps without participants necessarily needing a conceptual understanding that they serve this function, one might anticipate that hint versions would be easier than non-hint versions to solve. On the other hand, adding the blank cards might be seen to create a different task load, with four types of cards to place rather than three.

In the current experiment, participants were presented with both the hint version and the non-hint version of the Cards problem. We would expect a high level of transfer between the two conditions – the problems are superficially and conceptually very similar. Thus, there should be no difficulty in recognising the relevance of the problems to each other. If analogical transfer is supported by successful solution of the source problem, then the use of hints at time one (T1) should scaffold the problem-solver to a solution, effectively showing them the answer – we would expect a high rate of solutions and for solutions to be found quickly at time two (T2). An alternative hypothesis is that the use of hints in this way leads to a superficial understanding of the source problem and frustrates encoding of its conceptual structure even if it allows solution. If this is the case, we would expect solution rates to the non-hint Cards problem to be similar at T1 and T2.

Instructions check as a measure of conceptual understanding.

Experiments on problem-solving come with often unexamined assumptions. These assumptions can restrict the interpretation of participants' behaviour. One such aspect is the importance of the instructions in structuring problem solution. In problem solving research, the instructions serve the purpose of clearly explaining the parameters of the problem. Unless research explicitly examines the role of hints on performance (e.g., Ormerod, MacGregor & Chronicle, 2002), it is rare for instructions to offer advice on how to solve the problem because participants' ability to generate methods of problem solution forms part of the measured outcomes. Instructions, therefore, typically lie outside of the parameters of the problem-solving experiment. Participants are given the instructions, asked if they understand the parameters and are then tasked with solving the problem. It rarely reported whether the instructions are left or removed.

Detailed qualitative analysis from Ross and Vallée-Tourangeau (2021) demonstrated that the relationship between problem solution and instructions is likely to be complex. They presented participants with the Socks problem[2] and found that, despite detailed piloting of instructions and regular verbal checks during the reading of the instructions by the researcher, participants still sought clarification as the problem-solution process was ongoing. The authors describe the instructions as becoming an artifact shaped during the problem-solving process as different parts of the instructions were probed.

A possible explanation for this inability to clearly encode the instructions draws on cognitive load theory (Sweller, 2011). Retaining instructions in working memory reduces the ability to apply oneself fully to the problem-solving process. The most cognitively parsimonious activity is to encode only those aspects of the instructions that are required to make progress in problem solution. As the participant does not know the problem solution, initial hunches as to the best way to make progress may lead to shallow encoding of the key aspects of the instructions.

Indeed, we know that adult individuals in problem-solving experiments often demonstrate limited planning ahead and are prone to act before thinking. For example, Ormerod et al. (2013) found that, in attempting the Nine-ball problem described above, a surprisingly large number of participants (35%) chose for their first weighing to put five balls on one side and four on the other of the balance scale, an attempt guaranteed to carry no useful information since it confounds number and weight. Ormerod et al explain this seemingly irrational performance as resulting from participants choosing to maximize perceived (but erroneous) progress towards the problem goal without fully understanding the pre-requisites of the initial problem.

A strategy of maximizing perceived progress in initial solving attempts is particularly catastrophic in the class of problems deemed insight problems, which are constructed in such a way that the most attractive way of solving the problem initially leads to a dead end. Thus, an initial reading of instructions may trigger misdirection, and failure may generate a belief that the initial reading was incomplete or misunderstood, or that the instructions contain further information that was not sampled fully in failed attempts to solve.

Encoding of instructions, therefore, offers a potentially variable level of understanding about the problem structure. Repeated views of the instructions may serve, in this context, as an indicator of the level of source problem comprehension, independent of the participant's achievement of solution to it. To examine the role of hints in source comprehension using instructions checks as a measure, we conducted an experiment in which participants solved two problems: either a no-hint version of the Cards problem as source followed by a hint version as target, or vice versa.

We therefore had the following hypotheses:

[2] If you have a drawer with brown socks and black socks mixed in a ratio of 4:5, how many would you have to pull out in order to guarantee a pair? The answer is 3. The ratio is misleading: after three socks, a pair would have to be made.

H1: Success will be higher at T1 in the hint condition
H2: Success in the no hint condition will be greater at T2 than T1.
H3: Instruction checks will predict success in all conditions.

Method

Participants

One hundred and twenty participants (M_{age} = 38.6 years, *SD* = 14.2 years) were recruited from prolific.co. and were paid £5 for participation. One participant's data were lost due to technical failure. There were 104 males, 109 females, and 6 non-binary/ not disclosed, with genders distributed approximately equally across conditions.

Procedure

Participants were invited to solve the card problem (see above) twice with problem presentation varying by the presence or absence of the blank cards. This yielded the following two conditions: Hint first (KQJB → KQJ) and No hint first (KQJ → KQJB). The problem was programmed in PsychoPy and presented online in Pavlovia.

Participants were given the following instructions (alternative instruction components for the non-hint variant are shown in square brackets):

> In front of you on the next page you will see twelve [sixteen] cards taken from a normal pack of playing cards - four Kings, four Queens and four Jacks [along with three blank cards that can serve as guides].
>
> Arrange all the cards in a grid [using the blank cards as guides] such that each row and each column in the grid that contains any cards has precisely one King, one Queen and one Jack, no more and no fewer.

The screen that followed presented the cards which needed to be moved to enact the solution. The cards were presented at the start of each trial for all participants on the left of the screen as three offset stacks. one for each suit. Movement of cards was done by clicking on a card and dragging it with the computer pointer to the desired position, where it snapped to grid. Once all cards were moved to the chosen positions, participants submitted their solution. If participants entered an incorrect solution, they were given feedback that their solution attempt was incorrect and were invited to attempt again. Participants were also able to reset to the start or to review the instructions again. If they selected the 'review' button, which they could do as many additional times as they desired, they were shown the following "recap" instructions:

> Arrange all the cards in a grid such that each row and each column in the grid that contains any cards contains precisely one of each of the Royal Suit cards, no more and no fewer. Click on this screen to hide it in order to carry on with the problem.

Participants were given 5 minutes at T1 to solve the first card problem before being asked to solve compound Remote Associates problems taken from Sio & Ormerod (2015) for 5 minutes. Finally, they were presented with the second card problem at T2 and were also given 5 minutes to solve.

Data were collected for correct solutions, latency to solution and instruction checks at T1 and T2.

Results

Hypothesis One. Hypothesis one tested whether hints at T1 increased success rates. Overall success rates were low at T1 and only increased at T2 for the Hint condition (see Table 1).

Table 1: Solution rates at T1 and T2 as a functions of condition.

	T1	T2
No Hint - Hint	30%	57%
Hint – No Hint	36%	35%

There was not a significant difference in solution rates at T1, β = -0.03, 95% CI [-0.12, 0.06], $t(107) = -0.74$, $p = .458$; Std. β = -0.07, 95% CI [-0.26, 0.12]. At T2, the difference in the conditions was statistically significant, β = 0.11,95% CI [0.02, 0.21], $t(107) = 2.44$, $p = .015$; Std. β = 0.23, 95% CI [0.04, 0.41])

Participants averaged a similar amount of time to a correct solution at T1 (*M* = 185.9s, *SD* = 70.9s in the hint condition and *M* = 185.7s, *SD* = 57.6s in the no hint condition). This difference was not significant, (β = -0.10, 95% CI [-12.4, 12.2], $t(107) = -0.02$, $p = .987$; Std. β = -1.52e-03, 95% CI [-0.19, 0.19]). They were faster at T2 in the no hint condition (*M* = 166.2s, *SD* = 61.3s) and even faster in the hint condition (M = 136.7s, *M* = 69.5s). This difference was significant, (β = -14.75, 95% CI [-27.2, -2.3], $t(107) = -2.35$, $p = .021$; Std. β = -0.22, 95% CI [-0.41, -0.03])

Thus, hypothesis one was not upheld. There were no significant differences between the two types of problem at T1 in either solution rates or latency to solution. Hints did not seem to support problem-solving. Significant differences only emerged after working through a source problem.

Hypothesis Two. Hypothesis two tested whether there was an effect of transfer of performance in the no-hint and hint conditions. A linear model with time as predictor showed that there was no significant difference in proportion correct in the no-hint condition between T1 and T2, (β = -0.02, 95% CI [-0.11, 0.06], $t(107) = -0.55$, $p = .586$; Std. β = -0.05, 95% CI [-0.24, 0.14]). The difference in latencies is statistically non-significant (β = 9.75, 95% CI

[-1.54, 21.04], $t(107) = 1.71$, $p = .090$; Std. $\beta = 0.16$, 95% CI [-0.03, 0.35]). In the hint condition, however there was a significance difference in success rates - The effect of group is statistically significant and positive ($\beta = 0.11$, 95% CI [0.01, 0.20], $t(107) = 2.23$, $p = .026$; Std. $\beta = 0.21$, 95% CI [0.03, 0.40]) and the latency to solution was also significantly faster, the effect of group is statistically significant and negative ($\beta = -24.60$, 95% CI [-37.78, -11.42], $t(107) = -3.66$, $p < .001$; Std. $\beta = -0.33$, 95% CI [-0.51, -0.15]).

Hypothesis Two was not supported: there was no benefit to experiencing a hint condition prior to being presented with the no-hint condition. However, an exploratory analysis does suggest that experiencing the no-hint condition did improve success rates in the hint condition.

Hypothesis Three. We wanted to assess whether transfer was dependent on conceptual encoding. To do this we used instructions checks as a proxy, the fewer checks being made the greater the conceptual understanding of the problem. We first assessed if success was dependent on successful encoding via fewer instructions checks across all conditions. At T1, fewer instruction checks was significantly associated with greater success, $\beta = -0.04$, 95% CI [-0.06, -0.02], $t(107) = -4.50$, $p < .001$; Std. $\beta = -0.40$, 95% CI [-0.57, -0.22]) and also at T2, $\beta = -0.04$, 95% CI [-0.06, -0.02], t(107) = -4.08, p < .001; Std. $\beta = -0.37$, 95% CI [-0.54, -0.19]). This supports hypothesis three. From this result, we can infer that checking instructions indicates a lack of understanding.

We conducted an exploratory analysis to assess how conceptual understanding, as measured by instruction checks, relates to transfer. We looked at two categories of participants: (1) those who got it correct at T1 and incorrect at T2 (2), and those who got it incorrect at T1 and correct at T2. These two categories of participants were those for whom transfer (of either success or failure) either did not, or could not, occur. The average instruction checks can be seen in Table 2.

Table 2: Instructions checks at T2 as a function of success at T1, success at T2 and the presence of hints at T2.

	Incorrect T2	Correct T2
Correct at T1		
No Hint at T2	6.38 (5.55)	0.16 (0.39)
Hint at T2	N/A	0.29 (0.47)
Incorrect at T1		
No Hint at T2	2.79 (3.97)	0.42 (0.79)
Hint at T2	4.42 (7.31)	0.88 (1.41)

We constructed a linear model for those who got it correct at T1 to assess whether success could be predicted by the presence of hints at T2 or the number of instructions rechecks. The use of a hint was not a significant contributor to the model, $\beta = 0.10$, 95% CI [-0.15, 0.36], $t(33) = 0.78$, p =.436; Std. $\beta = 0.11$, 95% CI [-0.17, 0.40], but the number of instruction rechecks was, $\beta = -0.07$, 95% CI [-0.11, -0.04], $t(33) = -3.97$, $p < .001$; Std. $\beta = -0.57$, 95% CI [-0.86, -0.29]). For those who got it incorrect at T1, both the number of instruction rechecks, $\beta = -0.03$, 95% CI [-0.05, -6.54e-03], $t(70) = -2.57$, $p = .010$; Std. $\beta = -0.28$, 95% CI [-0.50, -0.07]) and the presence of hints, $\beta = 0.26$, 95% CI [0.06, 0.47], $t(70) = 2.52$, $p = .012$; Std. $\beta = 0.28$, 95% CI [0.06, 0.50]) were statistically significant.

Discussion

When first presented with the Cards problem, participants performed approximately equally whether they were provided with a hint or not. Also, participants performed equally poorly when they solved a no-hint variant as the source problem as when they solved when they solved a no-hint variant as the target (i.e., having received the hint variant as the source). However, solution rates for participants who received the hint variant as the target problem having solved a no-hint variant as the source problem were significantly higher than solution rates for the hint variant as source problem.

There are alternative explanations for these patterns of results. It may, on the one hand, be an effect of problem difficulty. It is well known that analogical transfer between easy and hard problems is asymmetric, with a harder source problem facilitating analogical transfer to an analogical target (e.g., Reed, Ernst, & Banerji, 1974). However, our results suggest no overall difference in problem difficulty between hint and non-hint variants when solved as source problems. Instead, the analysis of the instructions rechecks suggests that experiencing the no-hint condition scaffolded participants into developing a stronger conceptual understanding of the source problem, reflected in the significantly lower number of instructions checks at T2. Further exploratory analyses on those who got the answer correct at the first time suggest that the knowledge of the problem that was carried over was an understanding of the task entailed in the problem, as evidenced by fewer instruction checks.

The results we present here have several implications. First, we suggest that scaffolding exercises that are used in analogical transfer training methods need to be carefully constructed to elicit conceptual understanding of the underlying task demands of a training problem, rather than simply procedural understanding of a putative solution process for that problem. Second, we suggest that we need to be careful about ascribing lack of analogical transfer to a failure in the process of analogy itself. Studies have shown that analogical transfer can be impaired when the analogous nature of source and target problems is not recognized (e.g., Ross, 1987). Others have shown how superficial similarities can suggest false analogies between conceptually different problems (e.g., Novick, 1988). Our results indicate a different problem: solving a source problem successfully may be insufficient to enable analogical transfer if that solution process does not yield a deep conceptual understanding of the source itself. This finding is consistent with research that shows how analogical transfer can be enhanced when participants receive multiple analogs as source problems (e.g., Kurtz & Loewenstein, 2007) or are presented with an abstract schema for the analogous source and target problems (e.g., Robins & Mayer, 1993).

A question arises as to what it is that participants are seeking when then repeatedly inspect the problem instructions. One possibility is that they simply do not understand what the task entails at all (i.e., the constraint to have exactly one card of each denomination in each row and column). Another possibility is that participants do in fact have a conceptual understanding of the task at hand and are re-inspecting the instructions in case they are prompted to consider a solution idea proffered in the instructions that they may have overlooked. While these eventualities might be the case for participants who re-inspected the instructions at T1 and failed to solve at TI, it cannot be the case for participants who inspected the instructions at T2 having successfully solved at TI. Instead, we suggest that the re-inspection of task instructions is indicative of participants lacking a conceptual understanding of the problem. This is particularly the case for participants who solved at T1 but were unable to solve at T2. A future study in which concurrent verbalisations of participants are collected may shed more light on what they seek from instruction checks.

As in most laboratory studies of insight and analogical transfer, the problems and methods used here create limitations on the generality of our findings. Whether instruction checks provide a valid measure of source comprehension in richer domains such as education or design problem-solving remains to be tested. One possibility is that the insight and non-insight task variants are seen by participants as being essentially the same problem, and that the 'hint' offered by the blank cards is not interpreted as such. We are currently exploring the role of source comprehension as evidenced by instruction checks in studies where the differences between source and target are greater.

The inclusion of hints means this is not straightforward analogical transfer, and it may be that, rather than being too similar, the problems were not similar enough in structure to expect transfer. This would explain similar solution rates in the no hint condition at T1 and T2. However, were this the case we would also expect similar solution rates in the hint condition at T1 and T2. The increase in solution rates suggests that there was a benefit to experiencing the no hint condition prior to the hint condition and the differences in instruction checks suggests that this benefit was that participants knew what to do. We acknowledge that knowing what to do with a problem is not a direct proxy for conceptual understanding but we submit that our data show that it a closer proxy than problem solution.

We chose to explore instruction checks in the first instance using closely related problems, what Markman and Gentner (1996) term literal similarities. Our reasoning was that, if source comprehension limits transfer between literal similarities, then its impact is likely to be even greater as the conceptual and/or superficial distances between source and target problems increase. We note a comment by one reviewer of the current paper who stated "From the analogy literature, the idea that source understanding is a prerequisite for transferring the base solution to the target is so axiomatically true that I fail to recall anybody who bothered to pursue an empirical demonstration". We suggest that source understanding cannot be assumed and may provide an important limit on analogical transfer.

References

Akin, Ö., & Akin, C. (1998). On the process of creativity in puzzles, inventions, and designs. *Automation in Construction*, *7*(2-3), 123-138.

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

Cunningham, J. B., & MacGregor, J. N. (2008). Training insightful problem solving: Effects of realistic and puzzle-like contexts. *Creativity Research Journal*, *20*(3), 291–296. https://doi.org/10.1080/10400410802278735

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, No. 270.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*(3), 306–355.

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

Gozli, D.G. (2017). Behaviour versus performance: The veiled commitment of experimental psychology. *Theory & Psychology*, *27*(6), 741–758.

Keane, M.T., Ledgeway, T. and Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, *18*(3), pp.387-438.

Kolodner, J. L. (1997). Educational implications of analogy: A view from case- based reasoning. *American Psychologist*, 52, 57-66.

Kurtz, K. J., & Loewenstein, J. (2007). Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one. *Memory & Cognition*, *35*, 334-341.

Markman, A.B. & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition*, *24*, 235-249.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 510–520.

Ormerod, T. C. (2023). Possible, yes, but useful? Why the search for possibilities is limited but can be enhanced by expertise. *Possibility Studies & Society*, *1*(1).

Ormerod, T.C., MacGregor, J.N. & Chronicle, E.P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 28, 791-799.

Ormerod, T. C., MacGregor, J. N., Chronicle, E. P., Dewald, A. D., & Chu, Y. (2013). Act first, think later: The presence and absence of inferential planning in problem solving. *Memory & Cognition*, *41*(7), 1096–1108.

Ormerod, T.C., and MacGregor, J.N. (2017). Enabling spontaneous analogy through heuristic change. *Cognitive Psychology,* 99, 1-16.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, 6, 436–450.

Reeves, L., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, *115*(3), 381-400.

Robins, S., & Mayer, R. E. (1993). Schema training in analogical reasoning. *Journal of Educational Psychology, 85*, 529–538.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 629–639.

Ross, W., & Vallée-Tourangeau, F. (2021). Rewilding cognition: Complex dynamics in open experimental systems. *Journal of Trial and Error*, *2*.

Sio, U. N. & Ormerod, T.C. (2015). Incubation and cueing effects in problem-solving: Set aside the difficult problems but focus on the easy ones. *Thinking & Reasoning*, *21*, 113-129.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (55, pp. 37-76). Academic Press.