Learning and interactivity in solving a transformation problem

**Lisa G. Guthrie · Frédéric Vallée-Tourangeau ·
Gaëlle Vallée-Tourangeau · Chelsea Howard**

Abstract Outside the psychologist's laboratory, thinking proceeds on the basis of a great deal of interaction with artefacts that are recruited to augment problem-solving skills. The role of interactivity in problem solving was investigated using a river-crossing problem. In Experiment 1A, participants completed the same problem twice, once in a low interactivity condition, and once in a high interactivity condition (with order counterbalanced across participants). Learning, as gauged in terms of latency to completion, was much more pronounced when the high interactivity condition was experienced second. When participants first completed the task in the high interactivity condition, transfer to the low interactivity condition during the second attempt was limited; Experiment 1B replicated this pattern of results. Participants thus showed greater facility to transfer their experience of completing the problem from a low to a high interactivity condition. Experiment 2 was designed to determine the amount of learning in a low and high interactivity condition; in this experiment participants completed the problem twice, but level of interactivity was manipulated between subjects. Learning was evident in both the low and high interactivity groups, but latency per move was significantly faster in the high interactivity group, in both presentations. So-called problem isomorphs instantiated in different task ecologies draw upon different skills and abilities; a distributed cognition analysis may provide a fruitful perspective on learning and transfer.

Keywords Problem solving · Interactivity · Learning · Epistemic actions · Distributed cognition

L. G. Guthrie (✉) · F. Vallée-Tourangeau (✉) ·
G. Vallée-Tourangeau · C. Howard
Department of Psychology, Kingston University, Kingston upon
Thames KT1 2EE, UK
e-mail: l.guthrie@kingston.ac.uk
e-mail: f.vallee-tourangeau@kingston.ac.uk

Introduction

Problems are encountered frequently in everyday activity, varying in complexity and occurring across a diverse array of settings. In solving these problems, or indeed making sense of situations, people interact with local resources, both cultural and material (Kirsh, 2009). Traditionally, problem solving has been cast and understood in terms of information-processing models of move selection in a clearly defined problem space (Newell & Simon, 1972) or more recently of the shifts in excitatory and inhibitory activation in layered networks of "knowledge elements" that result in the restructuring of a problem representation in working memory (Ohlsson, 2011, p. 105). Emphasis on mechanisms of information processing do not foreground the co-determination of an agent's representation of the problem and a problem's physical presentation wrought by interactivity (Kirsh, 2009; 2013).

Transformation problems have been the focus of research in cognitive psychology for the past 50 years. In these problems, a well-defined space connects an initial and a goal state. Legal moves are defined in terms of simple rules and enacted with simple operators. Participants must reach the goal state by transforming the initial state through a series of intermediate states. A well-studied class of transformation problems are river-crossing problems. In these problems, objects—people, animals, or things—must be carried from one "riverbank" to another on a "boat" but with a set of constraints on moves that can be selected to reach the goal. A common version involves three missionaries and three cannibals (Reed, Ernst, & Banerji, 1974); another involves three hobbits and three orcs (Thomas, 1974). In transporting all cannibals and missionaries from one bank to the other, cannibals must not outnumber missionaries on either bank. The boat can take at most two passengers, and at least one. The problem space is relatively narrow since illegal moves cannot produce blind alleys of any depth (Reed et al., 1974) and can be completed in 11 steps. In

different versions, problem difficulty is a function of the rules that constrain the number of objects that can be moved at any one time, which combinations of objects are allowed on the boat, and which combinations can be left on either bank. The number of objects and the rules that govern their transport map out a problem space that links the initial state with all objects on one side of the river to a goal state with all objects on the other side of the river. Cognitive psychologists have used this task as a window on problem solving, particularly planning, search, and move selection (Reed et al., 1974; Simon & Reed, 1976). Such river-crossing problems have been used as a testing platform for a number of process models of search and move selection, strongly influenced by developments in AI (Simon & Reed, 1976).

Greeno (1974) suggests that individuals learn from repeated attempts at completing the river crossing task, reflected primarily through a sounder appreciation of which move is correct in each state. Reed et al. (1974) also investigated the effects of experiencing this type of problem twice in a series of three experiments, examining transfer and learning using analogous problems (e.g., the river-crossing problem and the jealous husbands problem). They found that learning occurred with repetition of the same problem; however, transfer of knowledge between analogous problems was limited. Knowles and Delaney (2005) reported that with certain incentives, illegal moves could be reduced with repeated attempts.

Interactive problem solving

The river-crossing task involves moving people or things across a surface and as such foregrounds the importance of interacting with a physical model of the task. However, interactivity in the river-crossing problem has never been an explicit and systematic focus of investigation. The manner with which the river-crossing task has been implemented varies a great deal across studies. For example, Reed et al. (1974) used different types of coins to represent missionaries and cannibals. Jeffries, Polson, Razran, and Atwood (1977) developed a basic computer interface where participants typed in the objects they wanted to put in the boat on a given crossing. The interface accepted only legal moves and updated the simple representations (often with letters and numbers, such as "3M" for three missionaries) on either side of the riverbank. Participants kept on typing in their moves until they managed to transport all objects from one bank to the other. Knowles and Delaney (2005) designed a more realistic interface with icons representing travellers against a backdrop of a river with two banks and a boat. Participants selected moves by clicking on the travellers, which then appeared next to the boat on the screen. In all these instances participants were never offered a three-dimensional work surface on which objects transparently corresponding to the scenario protagonists are manipulated

and moved by hand. In contrast, developmental psychologists who worked with the river-crossing task, being less sanguine about "formal operations" presumably, have taken care to design rich interactive thinking environments with physical materials representing the boat, the river, and figurines corresponding to the cover story characters (e.g., Gholson, Dattel, Morgan, & Eymard, 1989).

A more explicit experimental focus on interactivity may unveil important aspects of problem-solving performance, aspects that may correspond more closely to problem-solving performance as observed outside the laboratory. For example, there is evidence that in other transformation problems interactivity substantially transformed problem-solving behavior. Vallée-Tourangeau, Euden, and Hearn (2011) reported that mental set is significantly reduced in Luchins's volume measurement problems when participants interact with an actual physical presentation of the problem. The manipulation of water jars created a dynamic problem presentation revealing solutions that were not simulated mentally. The selection of moves was guided and governed by the pragmatics of manipulating real objects in a wet environment to achieve a goal, and participants were less likely to persevere using a more complicated solution for the test problems. In a river-crossing task, interactivity may help participants work out the quality of different moves not by simulating their consequences mentally, but rather by simply completing the move and observing the results. Such moves then are "epistemic actions" (Kirsh & Maglio, 1994): moves that may not, in themselves, necessarily help narrow the gap with the goal state, but rather provide information on what to do next. Kirsh and Maglio demonstrated that it is faster and easier to physically rotate the tetrominoes in Tetris than to simulate their rotation mentally, leading to better and more efficient problem-solving behavior. Move selection in the river-crossing task can be opportunistic, although not necessarily mindless; rather the strategic consequences of a certain move can simply be observed. In a high interactivity context, planning need not take place "in the head"—moves may not be premeditated; rather the trajectory through the problem space is enacted through the moves (cf. Vallée-Tourangeau & Vallée-Tourangeau, 2014).

Thus, in a high interactivity environment, there may be less pressure on reasoners to simulate mentally a path to a goal state and move selection may not be dictated by a plan (cf. Suchman, 1987). Problem-solving performance could well be influenced by the ease with which reasoners can enact moves. In a context that favors interactivity, participants may produce more moves in solving the river-crossing problem, but do so more quickly than in a context in which implementing a move is slower and more costly in terms of mental planning effort.

Some have argued that as a result, high interactivity may retard the acquisition of a more abstract representation of the task and hence may not lead to the same degree of learning

(O'Hara & Payne, 1998; Svendsen, 1991). With a river-crossing problem, a low level of interactivity may force participants to think longer before selecting a move and may encourage the development of a sounder appreciation of the logical structure of the task, which then helps participants transfer their knowledge to a different presentation of the same or similar problems. These participants, once presented with the problem a second time, but in a high interactivity condition, may proceed to solve the problem much faster. In turn, solving a river-crossing problem first in a high interactivity condition may promote a more procedural appreciation of the task that might be bound to the exact physical characteristics of the reasoning context and hence transfer poorly when participants complete the problem a second time, but in a low interactivity condition. The goal of the present experiments was twofold: To determine the impact of interactivity on performance in the river-crossing problem and to determine the amount of learning across two presentations of the problem as a function of interactivity.

Experiment 1A

Experiment 1A examined performance in the river-crossing problem when presented with or without artefacts as an aid to solution. This was measured in terms of number of moves, latency to completion, and latency per move. In a high interactivity condition, the problem was presented with a board, a raft, and six figurines: Participants had to move the raft and the figurines across the board to register a move until they had moved all six figurines from one bank to the other. In a low interactivity version, the problem was described on a piece of paper and participants were asked to verbalize the moves they would make to reach the goal. They completed the problem twice, once with the high interactivity version and once with the low interactivity version; the order was counterbalanced across participants. Experiment 1A employed a mixed design with interactivity level as the repeated measures factor and order—low interactivity first, high interactivity first—as the between subjects factor. As moves can act as epistemic actions, we predicted that participants would produce more moves, would solve the problem more quickly, and that hence latency per move would be shorter in the high compared to the low interactivity condition. We also predicted that participants would complete the second presentation of the task more quickly than the first since they would be familiar with the procedure and may well exploit an episodic record of their trajectory to help them select better moves, and select them more quickly. However, the nature of the experience during the first crossing as a function of interactivity level could influence the amount of learning. On the basis of the arguments formulated in O'Hara and Payne (1998; see also Svedsen, 1991), low interactivity forces participants to plan

and contemplate moves and their consequences; the additional time and effort encourage more deliberation, and as a result participants are more likely to develop a sounder understanding of the problem and select fewer but better moves. When the problem is experienced a second time, this time in a high interactivity condition, performance improvements should be steep. In turn, experiencing the problem in a high interactivity condition first may reduce the investment in deliberative efforts, perhaps mitigating the development of a more abstract, hence transferable, representation of the problem: There should be little evidence of learning when the problem is experienced a second time in a low interactivity condition. In light of the results obtained, and at the recommendation of reviewers of a previous version of this manuscript, we replicated this initial experiment: We refer to the two versions as 1A (the original) and 1B (the replication).

Sample size

Reed et al. (1974) studied the effect of transfer between two problems with similar problem states, the Missionaries and Cannibals and the Jealous Husbands problems. The experimental design for their Experiment 2—the first experiment was inconclusive and the third addressing issues too dissimilar from the ones explored here—was a two-factor mixed design, with problem type and order as the factors. They recruited a sample of 54 participants, with 50 successful solvers, 25 in each problem condition. For Experiment 1A we recruited 63 participants; after removing participants who did not complete the task and outliers, we conducted our analysis on a sample of 48. Experiment 1B is a replication of Experiment 1A, following an a priori power analysis using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) to estimate the sample size required to obtain a similar effect. The observed $\eta_p^2 = .149$ for the 2 x 2 interaction effect on latency per move in Experiment 1A corresponded to a large effect size ($f = .42$, see Cohen, 1992), with a correlation between repeated measures of .016. Based on these estimates, the a priori power analysis indicated that a total sample size of 40 would be sufficient to detect a similar effect size. Given the sample depletion due to participants not completing one or both attempts, as well as the possibility of having to remove long latencies to control for skewness, we recruited a similar number of participants for 1B as we had for 1A.

Method

Participants

Experiment 1A Sixty-three university undergraduates participated in the experiment in return for course credits. The data for three participants were incomplete, therefore unsuitable for analysis. Of the remaining 60 participants, nine did not

complete the river-crossing problem and were excluded from further analyses. Following tests for skewness for the completion latencies, a further three participants were removed from the analysis. The final sample was composed of 48 participants (41 females, M_{age} = 21.3, SD = 5.0).

Experiment 1B Sixty-five university undergraduate and postgraduate students participated in the experiment in return for course credits. Twelve participants did not complete the task and were excluded from further analysis; the final sample comprised 53 participants (43 Females, Mage = 21.5, SD = 3.93).

Materials and procedure

Chickens and wolves were the protagonists in a river-crossing scenario. The objective was for the six animals to be transported from the left riverbank to the right one. The selection of a move had to comply with the constraints and rules of the problem. The same instruction sheet explaining the objective of the task and the rules of the problem was used for both conditions and could be read by the participants throughout the duration of the task. The instructions read:

> "Three wolves and three chickens on the left bank of a river seek to cross the river to the right bank. They have a boat that can carry only two animals at a time, but there must always be an animal on the boat for it to move. However if at any time the wolves outnumber the chickens on either bank the wolves will eat the chickens. Thus you cannot move the animal(s) in a manner that will result in the wolves outnumbering the chickens on either bank. The goal of the task is to move all the animals from the left bank to the right bank."

In the low interactivity version of the task, the researcher transcribed each move as verbalized by the participant onto a record sheet. The record sheet was a simple representation of the raft between the left and right banks of the river, with slots to record the nature and number of the animals on either side (which was denoted with a 'C' for chickens and 'W' for wolves; see left panel of Fig. 1); each page represented only one move. At any one time, participants could only inspect their

previous move as they dictated their next move to the experimenter. As soon as the next move was dictated, the sheet with the previous move was turned over. Thus participants could not inspect a historical record of previous moves. Illegal moves proposed by the participant were noted, but the experimenter did not transcribe the nature of the illegal move on the recording sheet. Rather, participants were invited to re-read the task instructions to discover why such a move was not allowed.

Legal moves were the moves made by the participant from the first move to the final move that met the constraints or rules as set out in the instructions sheet available to all participants throughout each attempt, whereas illegal moves were denoted as any moves that did not meet these constraints. The decision to include all violations of the rules as illegal moves was made in order to measure the total number of moves completed by the participant during the entirety of the attempt. Knowles and Delaney (2005) did not include violations of the rules which negated the movement of the boat if either empty or carrying more than two passengers on the grounds that participants may make errors in using the computer interface or through lack of understanding of the rules. In the experiments presented here there was no computer interface to negotiate. In addition, the rules and instructions were available in a printed format for all participants throughout both attempts; in fact, participants were actively encouraged to refer to the rules throughout the task.

The high interactivity version of the task involved the use of six plastic figurines, three wolves (9 cm x 7 cm x 2 cm) and three chickens (4 cm x 5 cm x 1.5 cm), one pop-stick raft (9 cm x 6 cm), and a painted board (60 cm x 45 cm) representing the river and banks (see right panel of Fig. 1). As the participants interacted with the artefacts, the experimenter recorded the moves, but this record was never shown to the participants; as with the low interactivity condition this ensured that participants could not review the problem-solving trajectory. An illegal move prompted the experimenter to instruct participants to move the raft and the animals back to the previous state and, as in the low interactivity condition, they were invited to re-read the instruction sheet to determine which moves were possible. In both conditions participants were given up to 15 min to complete the river-crossing problem. Participants were not asked to prioritize the number of moves made or the time in making moves, nor were they
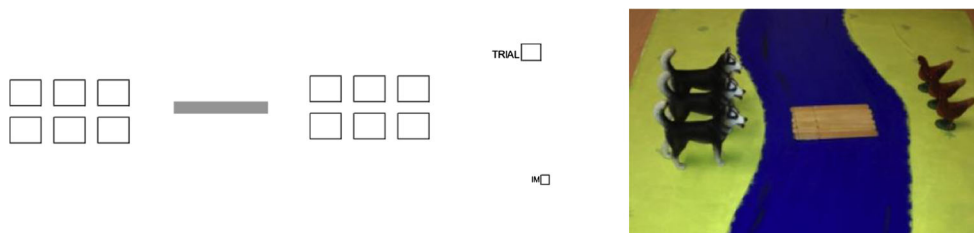

Fig. 1 Record sheet for the river-crossing moves in the low interactivity condition (*left panel*); board, raft, and figurines in the high interactivity condition (*right panel*)

explicitly told how long they would be given to complete the task. If the participant questioned the amount of time allowed to solve the problem, the researcher explained that a reasonable amount of time would be allowed within the confines of the experimental session time. However, any participant unable to finish one or both attempt within 15 min was excluded from subsequent analyses.

A 20-min interval was designed between the two presentations of the river-crossing problem during which participants completed a number of non-verbal puzzles, including finding similarities and differences between series of pictures, and identifying the odd picture in a series of thematically-related pictures. Finally, the river-crossing task was presented again in the alternate condition (either low or high interactivity) to that which was presented first; the order was counterbalanced across participants. Thus, the independent variables manipulated were condition (low interactivity, high interactivity) and order (low interactivity first, high interactivity first) in a 2×2 mixed design. Performance in both conditions was measured in terms of latency to solution, the total number of moves to solution, and latency per move. The latter offers the more interesting window onto problem-solving performance across interactivity conditions since it provides a gauge of how quickly, on average, participants generate each move. In keeping with previous river-crossing studies legal and illegal moves are reported separately. The latency per move data were calculated using the total number of moves.

Results

Latency

Experiment 1A Indices of skewness—as calculated following guidance in Fidell and Tabachnick (2003, p. 118)—indicated

that the latencies in three of the four experimental conditions were within the range of normality, but not in the low interactivity condition when experienced first. As mentioned above, removing three outliers in this condition resolved this problem ($Z = 1.2$). Latencies to solution, reported in Table 1, suggest that participants completed the second presentation of the task faster when they experienced the low interactivity condition first, followed by the high interactivity condition. A 2×2 mixed analysis of variance (ANOVA) revealed that the main effect of interactivity was not significant, $F(1, 46) = .606$, $p = .440$, $\eta_p^2 = .013$, while the main effect of order was significant $F(1, 46) = 8.17$, $p = .006$, $\eta_p^2 = .151$; the interactivity condition by order interaction was not significant $F(1, 46) = 2.70$, $p = .107$, $\eta_p^2 = .055$.

Experiment 1B Indices of skewness indicate that the latencies in the four experimental conditions were normally distributed. Latencies to solution are shown in Table 1; the pattern of findings closely replicated what was observed in Experiment 1A. The faster change in crossing latency was observed in the high interactivity condition when participants first completed the task in the low interactivity condition. A 2x2 mixed ANOVA showed the main effect of interactivity was not significant, $F(1, 51) = 3.45$, $p = .069$, $\eta_p^2 = .063$, while the main effect of order was significant, $F(1, 51) = 5.12$, $p = .028$, $\eta_p^2 = .091$; the interactivity condition by order interaction was also significant, $F(1, 51) = 9.76$, $p = .003$, $\eta_p^2 = .161$. Post hoc tests indicated that latencies in the low interactivity condition did not decrease significantly from the first to the second presentation, $t(51) = 0.358$, $p = .419$. In turn participants were faster in the second attempt at the problem than the first in the high interactivity condition, $t(51) = -4.097$, $p < .001$. When participants completed the low interactivity condition followed by the high interactivity condition, they were significantly faster in the second attempt, $t(23) = 4.297$, $p < .001$. When participants completed

Table 1 Mean latencies and mean number of moves to completion (along with their SD) in the river-crossing problem for each of the three experiments. order indicates the order of interactivity undertaken in the experimental session

		Latency (s)				Moves												
						Legal				Illegal				Total			
		First		Second		First		Second		First		Second		First		Second	
Experiment	Order	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1A	L/H	337.30	106.61	255.39	168.06	12.70	2.44	17.00	7.75	2.83	2.27	3.52	4.64	15.52	3.36	20.52	10.06
	H/L	426.24	189.48	397.00	234.63	15.84	4.96	13.88	4.25	5.72	4.42	3.64	3.41	21.56	8.45	17.52	6.81
1B	L/H	463.92	236.71	246.04	139.77	17.46	7.27	17.17	8.18	3.04	4.41	2.50	4.19	20.50	10.42	19.67	10.51
	H/L	493.97	267.51	438.62	271.15	19.07	7.22	16.97	7.66	7.55	4.61	4.14	6.87	26.62	10.14	21.10	13.29
2	L/L	408.12	165.04	217.21	78.35	13.45	3.80	12.06	1.87	3.39	2.76	1.15	1.66	16.85	5.65	13.21	2.47
	H/H	399.16	186.25	233.59	96.25	16.41	5.81	15.41	4.99	4.73	2.85	2.86	2.54	21.14	7.58	18.19	6.35

First and second represents the first or second attempt in the experimental session

L low interactivity, *H* high interactivity

the high interactivity condition first then the low interactivity condition there was no significant decrease in the time taken to complete the problem, $t(28) = .820$, $p = .419$.

Moves

Experiment 1A The high interactivity condition elicited a greater number of legal moves in solving the river-crossing problem compared to the low interactivity condition in both orders (see Table 1). In turn, the mean number of illegal moves was greater in the high interactivity condition than in the low interactivity condition when it was experienced first, but the frequency of illegal moves was relatively stable in the second presentation for both conditions. Thus, combining legal and illegal moves the total number of moves was always greater in the high interactivity condition. In a 2x2 mixed ANOVA for total moves the main effect of interactivity was significant, $F(1, 46) = 13.95$, $p = .001$, $\eta_p^2 = .233$, while the main effect of order and the interactivity by order interaction were not, $Fs < 1$.

Experiment 1B The high interactivity condition once again elicited a greater mean number of legal moves compared to the low interactivity condition in the first attempt (see Table 1). However, unlike Experiment 1A the number of legal moves in the second attempt were similar for both conditions. In turn, the mean number of illegal moves was higher in the high interactivity condition than the low interactivity condition when it was experienced first, but in the second attempt the number of illegal moves was lower in the high interactivity than in the low interactivity condition. Overall, then, total moves were greatest in the high interactivity condition for the first attempt but in a 2x2 mixed ANOVA the main effects of interactivity, $F(1, 51) = 1.27$, $p = .265$, $\eta_p^2 = .024$, and order, $F(1, 51) = 2.70$, $p = .107$, $\eta_p^2 = .050$, were not significant, nor was the interaction, $F(1, 51) = 2.34$, $p = .132$, $\eta_p^2 = .044$.

Latency per move

Experiment 1A The latency per move data—latency to completion divided by the sum of legal and illegal moves—are shown in the left panel in Fig. 2. Latency per move in the low interactivity condition was unaffected by order; however, participants appeared faster at enacting moves in the high interactivity condition, especially the second time the participants engaged with the task. In a 2×2 mixed ANOVA the main effects of interactivity, $F(1, 46) = 15.7$, $p < .001$, $\eta_p^2 = .255$, and order, $F(1, 46) = 8.90$, $p = .005$, $\eta_p^2 = .162$, were significant, as was the interactivity condition by order interaction, $F(1, 46) = 8.05$, $p = .007$, $\eta_p^2 = .149$. Post hoc tests indicated that the low interactivity latency per move was not significantly affected by order, $t(46) = -.105$, $p = .916$, whereas completing the high interactivity condition in the second attempt led to significantly faster move selection than in the first attempt in the high interactivity condition, $t(46) = -4.94$, $p < .001$. Further, participants were faster in making moves in the high interactivity condition following experience in the low interactivity condition, $t(24) = 8.32$, $p < .001$. Finally, there was no significant difference in how quickly moves were made in the low interactivity condition when experienced after the high interactivity condition, $t(22) = 0.641$, $p = .528$.

Experiment 1B The latency per total move data in the replication experiment are shown in the middle panel of Fig. 2. The pattern from Experiment 1A was replicated: Latency per move in the low interactivity condition appeared to be largely unaffected by order; however, participants were faster at enacting moves in the high interactivity condition, and this was particularly evident on the second attempt. In a 2x2 mixed ANOVA the main effect of interactivity was significant, $F(1, 51) = 39.17$, $p = .001$, $\eta_p^2 = .434$, but the main effect of order was not, $F(1, 51) = 1.42$, $p = .238$, $\eta_p^2 = .027$; however, the
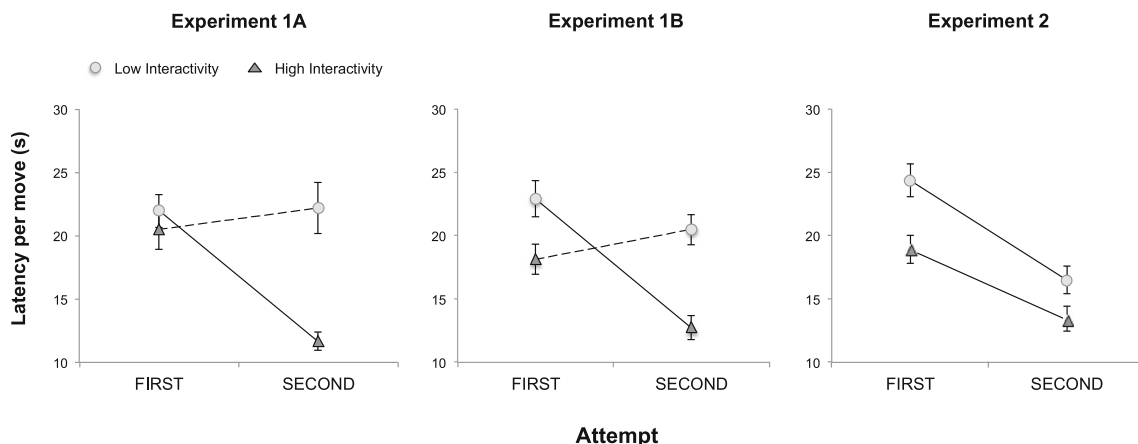

Fig. 2 Mean latency per move in the two interactivity conditions (*low, high*) for the first and second attempt for each of the three experiments. For the *left* and *middle panels*, Experiments 1A and 1B, the group that experienced the high interactivity condition first is represented by a *broken line* and the group that experienced the low interactivity condition first is represented by a *solid line*. Error bars are standard errors of the mean

condition by order interaction was significant, $F(1, 51) = 10.2$, $p = .002$, $\eta_p^2 = .167$. Post hoc tests indicated that the difference in latency per move was not significantly different between the two low interactivity attempts, $t(51) = .646$, $p = .521$, but the time taken to select a move during the second attempt in the high interactivity condition was significantly faster than during the first attempt, $t(51) = -3.42$, $p = .001$. In addition, participants were faster in making moves in the high interactivity condition following experience in the low interactivity condition, $t(23) = 8.36$, $p = .001$; latency per move remained unchanged when the high interactivity preceded the low interactivity condition, $t(28) = 1.97$, $p = .059$.

Discussion

Experiments 1A and 1B investigated the impact of interactivity on problem-solving performance for a river-crossing problem. In both experiments participants were required to solve the problem twice, once in a low interactivity context in which moves were simulated mentally and dictated to an experimenter and once in a high interactivity context where moves could be enacted through a three-dimensional display that corresponded to the main features and protagonists of the problem. The experiment was conducted twice to determine the robustness of the findings. In both the original experiment (1A) and its replication (1B) a high level of interactivity generally encouraged participants to make more moves in reaching a solution than when they completed the low interactivity condition. The latency per move data indicated that participants were always quicker to select a move in the high interactivity condition, and were generally quicker to select a move during the second presentation of the problem. However, the more important pattern in these data was the interactivity condition by order interaction observed in Experiment 1A and replicated in Experiment 1B: latency per move dropped precipitously when the second presentation of the problem occurred in the high interactivity condition. Notably, the post hoc tests on the latency per move data motivated by the significant order by condition interaction, revealed an identical pattern of results in both Experiment 1A and Experiment 1B. Thus, in both experiments order did not affect latency per move in the low interactivity condition, but did so in the high interactivity condition. In addition, there was no difference in latency per move between the low and high interactivity conditions if the high interactivity condition was attempted first, but latency per move was substantially reduced when the high interactivity condition was attempted second.

The replication of the latency per move data is particularly informative. However, it is important to consider all the data patterns including those not replicated. While the replication produced similar outcomes for six of the nine main results, to better evaluate the outcomes of the replication we take the three incongruous results in turn. First, the overall latency data showed a significant interactivity condition by order interaction in Experiment 1B but not in Experiment 1A. This may be explained by the fact that in 1A the mean latency recorded in the low interactivity condition when it was presented first was much lower than in 1B. However, the general trend was the same for both experiments with latency decreasing as would be expected from the first to the second attempt in both orders of presentation. Second, in the total moves data, the interactivity main effect was significant in 1A but not in 1B. This seems to reflect the greater number of moves during the first attempt in the low interactivity condition of Experiment 1B. As Table 1 shows, however, the difference in the mean total moves between the first attempt in the low interactivity condition and the first attempt in the high interactivity condition was similar at six moves for both experiments. Third, the latency per move data revealed a significant main effect for order in 1A but not in 1B. The important pattern in the latency per move data is the significant interaction, clearly illustrated in Fig. 2; however, in 1A, the mean latency per move was similar in both interactivity conditions during the first attempt, and, as a result, the drop in move selection latency in the high interactivity condition during the second attempt produced an overall order main effect (namely, the order Low first - High second led to lower move selection latencies than the order High first - Low second).

The second presentation of the problem offered the opportunity to gauge the degree of learning and transfer from one interactivity context to another. There was much evidence of learning, when the second opportunity to solve the problem took place in a context that favored a high level of interactivity: participants completed the problem in less time and selected moves at a faster rate than when the second presentation of the problem was in the low interactivity condition. In fact, when the low interactivity condition was experienced second, performance reflected little learning and transfer. This pattern of results suggests two competing explanations: (1) the process and quality of knowledge acquisition is different as a function of the level of interactivity or (2) interactivity is a performance facilitator and a high level of interactivity more clearly showcases learning. We evaluate each explanation in turn.

First exposure to the problem without much interactivity might have fostered the acquisition of a sounder and more actionable representation of the task and appreciation of an efficient sequence of moves to solution. In contrast, experiencing the problem in a context that fosters a high level of interactivity might not be accompanied by the same investment in cognitive effort, proceeding primarily on the basis of procedural learning, which in turn might interfere with the development of an accessible and transferable conceptual representation of the problem. As a result, when the problem is encountered for the second time in a condition without much

interactivity, the procedural knowledge does not facilitate transfer; however, when the second presentation occurs in the high interactivity condition, performance substantially benefits from the knowledge acquired on the basis of the experience in the low interactivity condition. The pattern of results from Experiments 1A and 1B offers some support for this conjecture.

Alternatively, the substantial improvement in the high interactivity condition when participants are presented with the problem a second time might not reflect differences in the type and quality of learning but rather release from a performance bottleneck. In other words, interactivity is a performance facilitator. Cognitive efforts and task demands are more exacting with low interactivity—as evidenced by the significantly longer latency per move. When participants encounter the problem a second time but this time can benefit from cheaper move selection by moving artefacts on the board, they experience a release from the cognitive demands of the low interactivity condition and are quicker at producing moves, and hence quicker to reach a solution. The river-crossing problem is a narrow analytic problem with a tightly defined problem space: participants can't wander off-piste very far. Faster move selection and the production of more moves leads to the goal state quicker.

The design of Experiment 1A and the replication 1B, however, cannot provide data to adjudicate the relative merits of these conjectures since attempt and interactivity level were not decoupled. Experiment 2 was designed to disentangle the effects of order and interactivity on learning by manipulating the levels of interactivity between subjects. Thus, as in Experiments 1A and 1B, participants completed the river-crossing problem twice; however, unlike in those two experiments, they did so either in a low or a high interactivity context both times. In this manner, Experiment 2 could provide data to determine the magnitude of learning as reflected in improvement in performance across the two presentations. In light of O'Hara and Payne's (1998) conjecture that planning impacts performance for different levels of interactivity, Experiment 2 also sought to measure independently participants' planning abilities by having them complete a series of Tower of London problems during the experimental session. The Tower of London (ToL) is a transformation problem used to assess planning skills in healthy and neuropsychological patients (Shallice, 1982; Ward & Allport, 1997; Unterrainer, Rahm, Leonhart, Ruff, & Halsband, 2003). We sought to determine the degree to which planning performance with a series of three-disk ToL problems could predict river-crossing performance. Specifically, we predicted that planning skills would correlate with river-crossing performance in the low interactivity condition; however, in a high interactivity context, the ease of selecting and implementing moves should level off individual differences in planning abilities.

Experiment 2

Method

Participants

Eighty-nine university undergraduates participated in exchange for course credits. Thirteen participants did not complete the river-crossing problem within the allocated time and were subsequently excluded from further analysis. Following tests for skewness in the latency data a further six participants were removed from the analysis to ensure the data were normally distributed. The final sample was composed of 70 participants (58 females, $M_{age} = 22.9$, $SD = 5.5$).

Materials and procedure

The same procedure designed for Experiments 1A and 1B was employed in Experiment 2 save for two changes. First, participants were randomly allocated to either the low or high interactivity group, and therefore completed the river crossing in the same interactivity condition twice. Second, participants completed a series of Tower of London (ToL) problems during the experimental session: half of the participants completed the ToL before the first attempt at completing the task, the other half after the second attempt. The ToL task was adapted from Shallice's (1982) version by using paper pegs printed on an A4 card and colored paper disks that could be moved about on the card. The rules were printed on an A4 sheet of white paper and read: (1) move only one disk at a time; (2) move only the top disk; (3) no more than two disks on the middle peg and no more than one disk on the shortest peg. The disks and pegs were placed in front of the participant in the initial state. The nature of the task was explained to the participant. It was also explained that the researcher would note the time taken and number of moves completed for each of the four problems. They were not given the opportunity to practice; no instructions were given on planning or time allowed for completion of the problems. Participants were then asked to read the rules before beginning the task, and to state that they understood the rules and the goal of the task. The rules were then removed from sight. Participants were shown four different goal-state configurations of the three disks on three pegs of different heights in an identical format to the card. At the start of each problem, the disks were set at the same initial state, participants were shown a new goal-state configuration and were then required to rearrange the disks one by one to match the goal state. Each problem could be completed in a minimum of five moves.

Results

Latency

The latency data were skewed in three of the four experimental conditions. Removing the six slowest participants ensured that indices of skewness were within the acceptable range (Z in all conditions < 1.96). The mean latencies to completion in both groups for both attempts are reported in Table 1. Latency to completion declined considerably from the first to the second attempt in both interactivity groups. In a 2x2 mixed ANOVA the main effect of attempt was significant $F(1, 68) = 63.7$, $p < .001$, $\eta_p^2 = .483$; however, the main effect of group was not , $F < 1$, nor was the attempt by group interaction, $F < 1$.

Moves

Participants in the high interactivity group produced a greater number of legal moves than those in the low interactivity group during the first and second attempt, but participants in both groups selected fewer moves during the second attempt (see Table 1). The mean number of illegal moves was also greater in the high interactivity group than in the low interactivity group (see Table 1); however, both groups selected fewer illegal moves during their second attempt. A similar pattern was produced with the overall number of moves: more overall moves were made in the high interactivity group than in the low interactivity group, but fewer moves were made in the second attempt and this in both groups. In a 2x2 mixed ANOVA on total moves the main effect of attempt was significant $F(1, 68) = 11.0$, $p = .001$, $\eta_p^2 = .139$, the main effect of group was significant $F(1, 68) = 21.3$, $p < .001$, $\eta_p^2 = .238$, but the attempt by group interaction was not, $F < 1$.

Latency per move

The latency per move data are illustrated in the right panel of Fig. 2. There are two patterns of note. First, latency per move decreased during the second attempt in both groups; second, high interactivity participants were faster at selecting moves than the low interactivity participants during both attempts. In a 2x2 mixed ANOVA the main effect of attempt was significant, $F(1, 68) = 77.9$, p < .001, $\eta_p^2 =. 534$, as was the main effect of group $F(1, 68) = 12.6$, p =.001, $\eta_p^2 =. 157$; the group by attempt interaction was not significant $F(1, 68) = 2.15$, p =.147, $\eta_p^2 =.031$.

Planning skills as predictor of performance

The mean latency to complete each of the four ToL problems was 26.9s ($SD = 13.5$) for participants in the low interactivity group, and 31.0s ($SD = 18.9$) for those in the high interactivity group; the difference between groups was not significant, $t(68) = -1.04$, $p = .302$. The mean number of moves for each ToL problem in the low interactivity group was 8.67 ($SD = 2.63$) and 9.94 ($SD = 3.71$) in the high interactivity group: the difference was not significant, $t(68) = -1.64$, $p = .105$. Thus, planning skills as determined by this measure did not differ between the two groups of participants. However, ToL latencies moderately correlated with the latencies for the first attempt, $r(31) = .344$, $p = .05$, and strongly with the second attempt, $r(31) = .524$, $p = .002$, in the low interactivity condition. Thus, the faster participants were at completing the ToL problems, the faster they were at completing the river crossing problems for both attempts in the low interactivity group. In contrast, participants' ToL latencies did not predict the time to complete the river-crossing problem in the high interactivity group, either for the first attempt, $r(35) = .071$, $p = .675$, or for the second attempt, $r(35) = -.151$, $p = .372$.

Discussion

Experiment 2 was designed to offer additional data to adjudicate the conjectures formulated in explaining performance during the second presentation of the task in Experiments 1A and 1B. In those experiments, latency and latency per move were substantially lower, suggesting significant learning, when the high interactivity but not the low interactivity condition was experienced second. Since interactivity was manipulated within subjects in these experiments, the nature of the learning experience during the first crossing was not controlled across the two presentations of the problem. A couple of conjectures were offered that could only be tested with an experiment where the interactivity level was manipulated between subjects, hence controlling for interactivity level across the two presentations.

The data from Experiment 2 painted a relatively unambiguous picture: learning was evident in both the low and the high interactivity conditions. Thus, the substantial learning in terms of reduced latency and latency per move in the low interactivity condition in this experiment lends some support to the notion that, in Experiments 1A and 1B, experience in a high interactivity condition first may have retarded transfer when the second attempt took place in the low interactivity context. In addition, this performance improvement cannot be attributed to a performance bottleneck caused by the relative cost of move selection in the low interactivity condition. In Experiment 2, participants who completed the first crossing in the low interactivity condition were significantly faster completing the second crossing in the same condition.

However, latency per move was faster in the high interactivity condition than in the low interactivity condition for both attempts at solving the problem. The absence of a significant interaction between interactivity and attempt in these data indicates that the performance improvement was similar in both

interactivity conditions. Indeed, the average decrease in latency per move from the first to the second attempt at completing the problem was 7.86s (SD = 6.78) in the low interactivity group, and 5.62s (SD = 6.06) in the high interactivity group, a non-significant difference, t(68) = 1.47, p = .147.

We conjectured that learning in the high interactivity condition of Experiments 1A and 1B when experienced during the second presentation of the problem reflected a sounder appreciation of the task obtained through a more sustained planning effort during the first presentation with low interactivity. This hypothesis is not supported by the data obtained in the present experiment. Researchers such as O'Hara and Payne (1998; see also Svendsen, 1991) have argued that a low interactivity environment—in which it is relatively more costly in terms of efforts to produce a move— encourages people to plan more before selecting a move which results in a richer and more transferable problem representation. However, in the experiment reported here, participants who completed the problem in the high interactivity condition did so significantly quicker during the second presentation. Admittedly, a lower cost structure, where moves are cheap, encourages more moves; indeed participants in the high interactivity group always selected more moves than participants in the low interactivity group. However, latency per move was significantly faster in the second presentation, and indeed significantly faster than the latency per move for participants in the low interactivity group. Thus, the argument that learning is mitigated by the low cost structure is not supported by the latency per move data reported here.

Finally, participants in both groups did not differ in their planning abilities as reflected by the similar average latency to complete each of the four ToL problems. Of greater interest was the fact that ToL latencies were strongly positively correlated with latencies to complete the river crossing task, for both attempts, in the low interactivity group. In contrast, the river-crossing latencies did not predict participants' performance on the ToL problems in the high interactivity group. To be sure, this pattern of correlational evidence indicates that higher interactivity may reduce the contribution of forward planning in the selection of moves. Perhaps of more importance, it suggests that high levels of interactivity may elevate the performance of participants with poor planning abilities, such that their performance is comparable to participants with higher planning abilities.

General discussion

Outside the psychologist's laboratory, scientists and lay people alike naturally create and build artefacts or recruit existing ones to configure highly interactive contexts of reasoning and facilitate problem solving. Thus, solving jigsaw puzzles involves physically juxtaposing different pieces to gauge their fit; in Scrabble, letter tiles are physically re-arranged to facilitate word production; in Tetris, tetrominoes are physically rotated to determine their optimal slot. And beyond puzzles and games, experts structure an external environment to support thinking. Scientists use physical objects and their arrangement in space to formulate and test hypotheses: Watson (1968, pp. 152-155) describes how he cleared his desk, cut out shapes corresponding to the four nucleobases, and manipulated them until he saw which ones could be paired to hold the double helix together (see Vallée-Tourangeau, under review).

The key driver of thinking in these examples is interactivity, how features of the world that configure a certain problem are arranged and re-arranged dynamically over time to evince a certain solution, to produce a desirable goal state. The work reported here shares a number of theoretical commitments with the seminal characterization of the role of external representations in problem solving outlined in Zhang and Norman (1994) as well as the elegant experiments and modelling efforts on soft constraints that determine when the world or internal memory is consulted in a wide range of tasks (e.g., Fu, 2011; Gray, Sims, Fu, & Schoelles, 2006). And while Zhang and Norman demonstrate how problem-solving performance is facilitated when certain rules and dimensions are externally represented, they say little about interactivity as such, although it is at the heart of their participants' performance. In turn, the interactive behavior described in the resource allocation experiments and models supporting the soft constraints hypothesis is one that reflects the quick iterative sampling of information from either an internal source— memory—and the world (Gray et al., 2006; Gray & Fu, 2004). These researchers do not address the role of interactivity in modifying the physical arrangements of a problem, and the contingent spatio-temporal itinerary that maps the problem's solution.

In the case of the river-crossing problem, interactivity as designed in the high interactivity condition does not change the nature of the problem or reconfigure it in a more cognitively congenial manner (unlike, for example, in Tetris or Scrabble; see Maglio, Matlock, Raphaely, Chernicky, & Kirsh, 1999). What interactivity does is promote a more fluid way to explore the problem space, involving as it were limited cognitive resources to enact changes. The state of the world— as modeled by the artefacts—cues the next move. What is particularly interesting then is the tight coupling between the agent and the world. The raft, animal figurines, and river board are better thought of as configuring a world that is a representative of the "real" world, not a representation of it (see Noë, 2012): participants in the high interactivity condition directly manipulate the world not unlike how scientists manipulate three-dimensional models (such as molecules; see Toon, 2011; Watson, above). This coupling may be maintained by perception-action loops that may not be mediated

by complex representations. In addition, this level of interactivity may be accompanied by a greater degree of task engagement. Svendsen (1991) reports that participants who experienced a greater degree of interactivity in a low implementation cost condition of the Tower of Hanoi enjoyed the task more, were more likely to recommend the interface, and believed it was quicker to use and solve the problem. The problem space that described the river-crossing problem employed in the low and high interactivity condition was the same, clearly. However, from an ecological perspective, the problems were different in the two conditions. The two presentations afforded a different behavioral repertoire, supported by different perceptual and cognitive processes. Thus, the cognitive ecosystem (cf. Hutchins, 2010) implemented in the low and high interactivity condition was different, and important questions about learning and transfer can and should be raised in these different ecosystems.

Previous work on the river-crossing problem demonstrated learning across repeated presentation of the same problem; however, evidence of learning transfer across analogous problems is more equivocal (Reed et al., 1974). Knowles and Delaney (2005), using computer-generated images, also investigated learning in the river-crossing problem by attempting to improve performance through the reduction of illegal moves. However, what was being learnt was not made clear beyond offering the conjecture that learning reflected "enhanced rule verification skills" (p.679). But these additional skills were not independently assessed and measured, and the conjecture does not offer much beyond a re-description of the data. Unlike Knowles and Delaney, the focus of the experiments reported here was not on improving performance in the number of moves made with a cost manipulation; rather it was to investigate how different levels of interactivity using artefacts, not computer-generated images, influenced problem-solving performance and learning.

The experiments presented here indicate that learning proceeds in both interactivity contexts. In addition, Experiments 1A and 1B offered a potentially interesting window on the nature of the transfer from low to high interactivity: participants performed the task fastest in a high interactivity context when they had first experienced it in a low interactivity context. Over repeated presentations of the same problem, we would expect performance to improve, as it did. However, a change in problem-solving mode may mitigate the learning effect in that participants must learn a new way to play the game as it were. The data reported here suggest that it is easier to adapt when the change is from a low to a high interactivity condition, rather than the reverse. The embodied immediacy of the contact with the problem, unmediated by symbolic representations, favors a quicker selection of moves, which results in a quicker solution of the problem. The transfer from a high interactivity to a low interactivity condition results in a slower adaptation to the change in the task ecology because

move selection is mediated through an indirect symbolic representation of the task. The results of the second experiment make clear that both interactivity conditions promote learning, although participants remain quicker in the high interactivity condition. However, adapting to a new task environment is more challenging when participants move from a high to a low interactivity condition. Problem-solving performance is more efficient if participants experience its abstract version first and then subsequently engage with it with a greater degree of interactivity.

The transfer asymmetry observed in Experiments 1A and 1B has interesting implications for an understanding of problem isomorphs especially as teaching and learning tools. Knowledge acquired through traditional abstract teaching methods followed by an interactive experience using manipulable artefacts may result in improved transfer and learning beyond the expected learning of repeated problem presentations. Lave (1988) in her ethnographic research investigated the transfer of knowledge learnt at school as it is applied to everyday life, revealing the different heuristics "just plain folk" (p. 43) employed in solving math problems in practical situations, such as shopping in the supermarket. Lave questions the common view that arithmetic as learned in school, is "carried away" from this supposedly "context-free" learning environment as a transportable tool for direct application to practical situations requiring math skills. She suggested that rather than conceptualizing the individual as a "self-contained, disembodied technology of cognition" it is more appropriate to analyse "the whole person in action, acting with the settings of that activity" and interacting with the lived in world (p.17). Lave was concerned with the transfer of knowledge from abstract traditional teachings in school to more practical use in everyday life. The experiments reported here may lay a foundation to address these concerns with empirical evidence indicating that learning an abstract concept followed by consolidation with concrete three-dimensional recognizable artefacts embodied with the same rules and constraints of the original concept significantly enhances learning.

O'Hara and Payne (1998) discussed the planning-acting continuum in analytic problem solving, and investigated the task, environment, operator, and agent contingencies that influence the amount of planning. Clearly, a high degree of interactivity may encourage quicker and more fluid action that shapes and reshapes the problem presentation without much prior planning. A lower degree of interactivity may force reasoners to think more carefully in the process of identifying the best move in a sequence. While a higher level of interactivity enhanced the performance of an individual in terms of how quickly moves were made, when measuring the number of moves taken to complete the problem, low interactivity encouraged better performance. Therefore, determining which level of interactivity better promotes learning can only be answered relative to considerations of efficiency themselves

relative to a particular context of reasoning. In other words, the cost structure for a particular task is relative to a set of situated parameters: sometimes it is useful to think long and hard (e.g., carefully planning a move in chess) and sometimes it is best to act quickly (e.g., moving a zoid in Tetris). The efficiency metric is dependent on the situation: if it is costly to make moves, then it is important to invest time into the contemplation of each move to be made; if the number of moves made is unimportant, but time is of the essence, then acting quickly is the efficient use of available resources. Still, there remains an important challenge for research on interactivity, namely, to determine the nature of the learning and the knowledge representation evinced by different levels and modes of interactivity.

In the experiments presented here, the problem does not change between problem presentations; however, interactivity in the physical world alters the cognitive landscape. In addition the amount of transfer and learning were seen to be contingent on the sequence with which the problem presentation was experienced. These results suggest that outcomes for learning and problem solving may differ when using the body and the concrete world to explore the problem space. Education is rapidly advancing toward the use of computer-centered learning; coupling bodies to a dynamic and modifiable world during learning, problem solving, and decision making poses important pedagogical questions (Kirsh, 1997; Klahr, Triona, & Williams, 2007; Moreno & Mayer, 2007; Renken & Nunez, 2013). Interactivity is now often couched in terms of interaction with a computer interface, knowledge and skills learnt from that mode of interactivity need to be assessed against interaction with the lived in physical world. Future research on this front would likely yield findings with important pedagogical implications as well as offering guidance to researchers working on the innovation and learnability of scaffolding interfaces (cf. Bolland, 2011).

References

Bolland, B. R. (2011). Distributed cognition as a lens to understand the effects of scaffolds: The role of transfer of responsibility. *Educational Psychology Review, 23,* 577–600.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160.

Fidell, L. S., & Tabachnick, B. G. (2003). Preparatory data analysis. In I. B. Weiner, J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology* (Research methods in psychology, Vol. 2, pp. 115–141). Hoboken: Wiley.

Fu, W.-T. (2011). A dynamic context model of interactive behavior. *Cognitive Science, 35,* 874–904.

Gholson, B., Dattel, A. R., Morgan, D., & Eymard, L. A. (1989). Problem solving, recall, and mapping relations, in isomorphic transfer and nonisomorphic transfer among preschoolers and elementary school children. *Child Development, 60,* 1172–1187.

Gray, W. D., & Fu, W.-T. (2004). Soft constraints behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science, 28,* 359–382.

Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113,* 461–482.

Greeno, J. G. (1974). Hobbits and orcs: Acquisition of a sequential concept. *Cognitive Psychology, 6,* 270–292.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science, 2,* 705–715.

Jeffries, R., Polson, P. G., Razran, L., & Atwood, M. E. (1977). A process model for Missionaries-Cannibals and other river-crossing problems. *Cognitive Psychology, 9,* 412–440.

Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Science, 25,* 79–96.

Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 264–306). Cambridge: Cambridge University Press.

Kirsh, D. (2013). Thinking with external representations. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, cognition and human artifice* (pp. 171–194). Dordrecht: Springer.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic actions. *Cognitive Science, 18,* 513–549.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching, 44,* 183–203.

Knowles, M. E., & Delaney, P. F. (2005). Lasting reductions in illegal moves following an increase in their cost: Evidence from river crossing problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 670–682.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge: Cambridge University Press.

Maglio, P. P., Matlock, T., Raphaely, D., Chernicky, B., & Kirsh, D. (1999). Interactive skills in Scrabble. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Conference of the Cognitive Science Society* (pp. 326–330). Mahwah: Lawrence Erlbaum Associates, Publishers.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19,* 309–326.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.

Noë, A. (2012). *Varieties of presence*. Cambridge: Harvard University Press.

O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology, 35,* 34–70.

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. New York: Cambridge University Press.

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

Renken, M. D., & Nunez, N. (2013). Computer simulations and clear observations do not guarantee conceptual understanding. *Learning and Instruction, 23,* 10–23.

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transactions of the Royal Society of London B, 298,* 199–209.

Simon, H. A., & Reed, S. K. (1976). Modelling strategy shifts in a problem-solving task. *Cognitive Psychology, 8,* 86–97.

Suchman, L. A. (1987). *Plans and situated actions: The problems of human machine communication.* Cambridge: Cambridge University Press.

Svendsen, G. B. (1991). The influence of interface style on problem solving. *International Journal of Man-Machine Studies, 35,* 379–397.

Thomas, J. C., Jr. (1974). An analysis of behavior in the Hobbits-Orcs problem. *Cognitive Psychology, 6,* 257–269.

Toon, A. (2011). Playing with molecules. *Studies in History and Philosophy of Science, 42,* 580–589.

Unterrainer, J. M., Rahm, B., Leonhart, R., Ruff, C. C., & Halsband, U. (2003). The Tower of London: The impact of instructions, cueing, and learning on planning abilities. *Cognitive Brain Research, 17,* 675–683.

Vallée-Tourangeau, F., Euden, G., & Hearn, V. (2011). Einstellung defused: Interactivity and mental set. *Quarterly Journal of Experimental Psychology, 64,* 1889–1895.

Vallée-Tourangeau, G., & Vallée-Tourangeau, F. (2014). The Spatio-temporal dynamics of systemic thinking. *Cybernetics and Human Knowing, 21,* 113–127.

Ward, G., & Allport, A. (1997). Planning and problem solving using the five disc Tower of London task. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 50,* 49–78.

Watson, J. D. (1968). *The double helix.* London: Penguin.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18,* 87–122.