Diagrams, Jars, and Matchsticks: A Systemicist's Toolkit

Frédéric Vallée-Tourangeau and Gaëlle Vallée-Tourangeau

Kingston University

Author Note

Correspondence concerning this article should be addressed to either Frédéric Vallée-Tourangeau or Gaëlle Vallée-Tourangeau, Department of Psychology, Kingston University, Kingston-upon-Thames. UNITED KINGDOM, KT1 2EE, f.vallee-tourangeau@kingston.ac.uk or g.vallee-tourangeau@kingston.ac.uk; tel: +44 (0)208 417 2000, fax: +44 (0)208 417 2388. We thank Ryan Tweney and two anonymous reviewers for thoughtful comments on a previous version of this manuscript.

Abstract

Participants in cognitive psychology experiments on reasoning and problem solving are commonly sequestered: Efforts are made to impoverish the physical context in which the problem is presented, decoupling people from the richer and modifiable environment that naturally instantiates it outside the lab. Sense-making activities are constrained, but this conforms to the strong internalist and individualist commitments implicit to these research efforts: Cognition reflects internal computations and the scientists' toils must focus on the individual and what she is thinking, decoupled from the world. We contrast this position with one that identifies cognition as the product of a cognitive system that is configured and enacted by, minimally, an agent and the world in which she is embedded. We review work on the psychology of hypothesis testing and problem solving and argue that refocusing research efforts on the dynamic agent-environment couplings that generate cognitive products—such as a problem representation, a hypothesis or a problem solution—offer a much richer set of methodological opportunities to unveil how people actually think outside the cognitive psychologist's laboratory. We conclude by exploring the ontological implications of a systemic perspective on cognition.

Diagrams, Jars, and Matchsticks: A Systemicist's Toolkit

Isolating an individual in an artificial laboratory environment to better examine the cognitive processes implicated in a carefully crafted task reflects a commitment to methodological sequestering (cf. Bolland, 2011). The sequestering is considered essential for the scientific conduct of the enterprise in the same way that much of biology and chemistry is conducted under laboratory conditions: to isolate, control, parse, and generally reduce the complexity of multiply-determined and dynamic phenomena to fit the scientist's measuring instruments and formal models. The resulting window onto cognition throws up data, analysed, scrutinized and replicated. These data advance our understanding of human cognition to the extent that the crafty task that elicited them captures some essential feature of human cognition once the participant is released from the laboratory and allowed to roam free in the wild, as it were. This methodological commitment is often coupled to another one, namely methodological individualism: The unit of analysis is the individual and performance elicited from the crafty laboratory task is a reflection of her internal capacities and abilities (cf. Malafouris, 2013, p. 25). In social psychology, attribution theorists speak of a fundamental attribution error: observers' tendency to attribute the cause of an actor's behavior to internal disposition and "underestimate the potential impact of relevant environmental forces and constraints" (Ross, Amabile, & Steinmetz, 1977, p. 492). These methodological commitments make it impossible—by the very nature of the science that is conducted under their auspices—to understand cognition from a systemic perspective, as reflecting a coalition of internal and external resources. It is exactly this coalitional perspective that will be explored in this paper, and the aim will be to illustrate how much richer, and arguably more representative, a characterization of human cognition is possible as a result.

The paper is structured in three sections. In the first we explore the psychology of hypothesis testing as illustrated on the basis of performance in a celebrated rule discover task, namely Wason's (1960) 2-4-6 task. The unflattering performance profiled with this task has generally puzzled psychologists, and most have sought to identify the cognitive and personality features of the minority who solve the task to shed light on why so few do. We demonstrate, in contrast, that a focus on the physical context of the problem, and the representational richness it proffers, casts a more productive—and kinder—light on people's hypothesis testing behaviour. In the next section, we review recent work on the psychology of mental set and insight that explored the role of interactivity with the physical components of a problem situation in fostering more creative problem solving behaviour. The paper closes by outlining the methodological and ontological implications of casting reasoning as the product of singular cognitive ecosystems.

Hypothesis Testing and Discovery

Evans (in press) reviewed the past 50 years of research on hypothesis testing and discovery driven by Wason's (1960) simple concept attainment task, the 2-4-6 task. In this task participants must discover a rule that governs how three numbers are put together: They do so by generating new sequences of three numbers for which they receive feedback. On the basis of this feedback participants gain information and test new sequences until they are sufficiently confident to announce the rule. To start things off, the experimenter offers a number sequence that conforms to the to-be-discovered rule: 2-4-6. The features of this initial example lure participants to contemplate hypotheses such as even numbers increasing by a constant or some specific arithmetic operation (e.g., such as the third number is the sum of the first two): yet the target rule is ‘any increasing sequence’. The initial triple suggests rules that are a subset of a more general rule; as long as the sequences go up, the feedback will be positive. A positive test strategy where participants offer positive instances of their current hypothesis—e.g., current hypothesis is ‘evens increasing by 2’, and the sequence offered is ‘10-12-14’—will result in positive feedback. However, in this task, positive feedback shores up the sufficiency of the hypothesis, it does not test its necessity. The abundance of positive feedback, coupled with a positive test strategy make the target rule very difficult to discover. As Wason (1960) and many subsequent replications demonstrated (e.g., Gale & Ball, 2006) only about 20% of the participants discover the ascending number rule.

Wason designed this task to determine the degree to which people would naturally seek to falsify their hypothesis, very much in the spirit of the kind of epistemology promoted by Popper at the time (the English translation of *Logik der Forschung* was published in 1959): In essence, to examine the willingness of reasoners to entertain counterfactual hypotheses and predict what could happen if their pet hypothesis was false. A more important question perhaps is whether the 2-4-6 task is representative of inductive reasoning and hypothesis testing, that is whether one can learn anything about inductive reasoning by using this task. As lab tasks go, its structure and content are far removed from scientific hypothesis testing, and probably from a wide range of more quotidian efforts. However, as in the world outside the lab, reasoners are asked to determine the scope and truth of a hypothesis, and in that way that task offers an interesting mean of determining how people rise to that inferential challenge. Some have argued that the pragmatics of the situation turns this into an unfair task (Vanderhenst, Rossi, & Schroyens, 2002): Why would the experimenter provide the initial triple ‘2-4-6’ if it was not informative? Yet, there are many instances in the history of science where the naïve interpretation of natural phenomena prevents the development of the correct explanation—heliocentrism for example. In other words, nature is indifferent to the untutored mind. But we’re straying: There are other features of the 2-4-6 task that make it less representative of hypothesis testing outside the laboratory, namely the restricted level of interactivity with artefacts and the resulting representational poverty of the triple sequences.

Rule discovery is a form of inductive inference, and as such, there exists no method that can guarantee the truth of an inference, however much adduced evidence appears to support it. In turn, participants may announce the correct rule on the basis of the flimsiest of evidence. Predictors of discovery in this task include (i) hard work, as measured by the number of sequences generated before announcing a rule, and (ii) creativity, as measured by the breadth of the number sequences generated, including descending sequences, sequences that go up in variable increments (e.g., ‘2-4-13’; Vallée-Tourangeau, Austin, & Rankin, 1995). Hard work and creativity help participants produce a more representative sample of number sequences, from which they are more likely to infer the correct rule. But there is no guarantee: the mechanics of inferential reasoning are laced with uncertainty.

Hard work and creativity seem like the perfect examples of individual, dispositional, factors. Some work harder, some are a more creative, and these individuals are more likely to succeed; ergo, a successful psychology of rule discovery must understand hard working and creative people. Wason’s initial characterization of the lucky 20% who solved the task was in terms of these people’s “disposition to refute” (1960, p. 139). The implicit call to arms with this conclusion is for psychologists to determine what individual differences—in terms of cognitive or psychological dispositions—could explain a participant’s ability to infer the correct rule. These dispositions vary across people, and hence some are more likely to succeed than others. Such dispositional conjectures, however, deflect researchers’ attention away from factors external to participants, that is away from the context—literally the physical context—of reasoning.

There is also considerably more explanatory upside for cognitive psychologists in exploring features of the context of reasoning, features that are external to the reasoner, than to keep the context of reasoning constant—i.e., to maintain the same experimental procedure—but hunt for the minority of participants who can solve the problem. On the one hand, keeping the context of reasoning constant implicitly endorses the representativeness of that context; using the same procedure across experimental demonstrations assumes that it provides a representative window onto the cognitive processes that are essential to perform the task. On the other hand, there is a form of cognitive elitism in letting individual differences determine cognitive performance on the task. The majority who can't crack the problem is cast aside as a relatively uninformative source of data, and the research efforts target the minority who outperforms.

Diagrams. A powerful way to elevate performance in the 2-4-6 task is to offer a diagrammatic representation of the number sequences as participants produce them. Vallée-Tourangeau and Payton (2008) designed a 2-4-6 task isomorph wherein participants produced new sequences by entering number in a bespoke Excel spreadsheet that simultaneously plotted the resulting pattern produced by the number sequence. As participants tested sequences such as '6-8-10' (Fig 1a.), '6-12-10' (Fig 1b.) or '1-5-13' (Fig 1c.), they also generated diagrammatic representations of the sequence.

Insert Figure 1 about here

Compared to a group who used a similar computer-based interface but one that did not synchronously produce a diagrammatic representation of the numbers, participants were twice more likely to announce the correct 'any ascending sequence' rule. Diagrammatic representations supported discovery by offering reasoners multiple representations of the same information, as well as the opportunity to switch from one medium to the other if stuck along an unproductive line of reasoning. Second, diagrammatic representations afforded perceptual inferences that may not be constrained by the arithmetic affordances of the initial triple '2-4-6' (e.g., its features suggest a formulaic progression). Indeed, participants who solved the task in the diagrammatic representation group were less likely to formulate algebraically specific hypotheses than participants who completed the task without the visual diagrams. And transcending the specific rules implicit to the initial triple is key: the diagrams encouraged more qualitative descriptions of the number sequences tested. Third, the interface afforded interactivity. That is as participants entered or deleted a number in the creation of a new triple, the graphical display would be updated synchronously. The immediate visual feedback cued actions and likely guided the subsequent selection of numbers. The agent and the interface are more tightly linked, configuring a dynamic system that produces new triples. This dynamic coupling transforms the agent and the nature of the triple generation process. The immediate and synchronous feedback may have favored triple production that emerged from unmediated action-perception loops. The hypothesis-testing narrative articulated by the participants as they formulate their next hypothesis may be the product rather than the cause of this triple generation process.

Thus, a modification of the context of reasoning substantially enhanced people's inductive inference abilities in a task that is otherwise very difficult, solved by a minority of participants. By deflecting attention from the nature of the participants to the nature of the context of reasoning, researchers are able to understand the circumstances in which people reason better. A better understanding of the context of reasoning alerts the social scientist, educator, ergonomist, among others, of the intervention opportunities to create a felicitous environment that supports sounder inferences.

Mental Set and Insight

The influential Gestalt psychologist Max Wertheimer (1959) offered a broad taxonomy of thinking into reproductive and productive types. The former reflects the application of well-rehearsed routine to the solution of problems. The latter is less anchored in past experiences, hence more ‘creative’, the kind of thinking that kindles genuine insight (insight is a state of knowledge, not a process; Dominowski & Dallob, 1995). Abraham Luchins, a Wertheimer student and one-time research assistant, developed and tested a procedure to measure the extent and impact of reproductive or ‘mechanized’ thinking with a simple volume measurement task involving three jars containing different volumes of liquid. In this task, participants are required to obtain an exact quantity of liquid, which does not correspond to the volume of any of the three jars (Luchins, 1942). Rather, they must engage in various pouring and discarding maneuvers to obtain the right amount. In the typical set up, participants are given a medium size jar (A), a large jar (B) and a small jar (C). Thus a problem might be: assume jar A contains 21 units, B 127 units, and C 3 units, obtain exactly 100 units of liquid. This would require filling up jar B (127), and from B pouring liquid to fill A (127 - 21 = 106), and from the remaining liquid in B, filling C once (106 – 3 = 103), discarding the liquid from C, and then repeating the maneuver, hence 106 – 2(3) = 100. In other words, the solution to this problem is B - A - 2C.

Luchins (1942) constructed a series of problem wherein the first five could be solved with the rule B – A – 2C; he called those training or *einstellung* problems (einstellung for latent expectation). Most participants learn the rule and when the fifth problem is presented, readily apply it. The next five problems are the interesting ones: they offer a window on the ‘mechanization of thought’ (the title of Luchins’s original monograph). Problems 6 and 7 can be solved with the B-A-2C rule, but can also be solved with a simpler A-C or A+C rule (e.g., obtain 20 units with A = 23, B = 49, and C = 3; the B-A-2C rule produces the right amount, but so does A-C). Participants who have not been trained with the first five problems, readily see the simple A-C solution; however participants who learned to solve the previous five problems continue using the more complicated algorithm. Problem 8 is particularly interesting: termed the ‘set breaker’, this problem can only be solved using the A-C rule, the complicated algorithm does not yield the right answer. Luchins’s spectacular finding was that the majority of his participants—and he tested a very wide variety of participants over countless replications of the effect—did not solve the set breaker problem although those who had not been exposed to the training problems could readily do so. Problems 9 and 10 were similar to the 6 and 7 in that they could be solved with the simple or more complicated algorithm. No matter their experience with the set breaker problem, participants continued using the B-A-2C rule for the last two problems.

Mental set, or the Einstellung effect, is a very robust phenomenon. A number of explanations have been put forward (e.g, Jensen, 1960; Bilalić, McLeod, & Gobet, 2008) essentially in terms of a schema acquired during training with the initial set of five problems, that is reliably activated by the unchanging configuration of the next five; the activated schema triggers the learned operators that are applied to solve the problem. The schema exerts control over the participants’ behavior.

The traditional methodological commitments to sequestering and individualism are particularly striking in the work on Einstellung: Luchins and those who studied the einstellung effect after him present this task to participants as a series of pen and paper arithmetic problems. The implied wetness in the description of the water jar task above in terms of filing, pouring and discarding maneuvers was strictly figurative: People never interacted with jars and water in the typical set up (there are interesting exceptions described in Vallée-Tourangeau, Euden, & Hearn, 2011). Contrast this with a situation where each of the water jar problems were presented in the same order—five training problems followed by five critical or test problems—but each presented to participants in terms of three actual jars of varying size next to a sink. In addition, participants are not provided with pen and paper, and instead must determine the target amount of liquid by literally performing filling and pouring maneuvers. For a start, without pen and paper the well-schooled arithmetic reflex is defused. The thinking ecology without pen and paper then may naturally encourage a different way of thinking about the problem. Certainly, the nature of the interaction with the artefacts is predicated on a completely different range or perceptual and motor information and possibilities. The schematic control over behavior during the test phase with the traditional procedure may be overcome by action-perception loops guided and constrained by the artefacts facing the participants. In the series of experiments reported in Vallée-Tourangeau et al. (2011), this is exactly what was observed. When participants were confronted with real jars at a sink, but with the identical series of Luchins's water jar problems, they readily acquired the complex pouring sequence during the training problems, corresponding to the B-A-2C rule, but were significantly more likely to opt for simpler pouring maneuvers for the test problems; they were also significantly more likely to solve problem 8—the extinction problem—than control participants who completed the task using pen and paper. Thus Luchins's celebrated window onto reproductive problem solving and the mechanization of thought is a procedural spandrel that reflects a deeply entrenched commitment to a model of thinking as disembodied abstract computation. In the wet version of Luchins's task, the system configured in terms of agent, jars, tap and sink, interlinked by action-perception loops, fosters much more efficient performance. If behaviour in the traditional version of the Luchins task is understood to reflect the influence of a problem solving schema acquired during the learning phase and triggered by the similarity with the test problems, control over behaviour in the interactive version is distributed among features of the entire system. Participants are in a better position to exploit efficiencies in the environment by directly recognizing the affordances of the concrete artefacts. The ecosystem instantiated with the real jars involves hand movements adjusted and guided by dynamic action-perception loops, and the basic perceptual features of the environment exerts a greater control over the agent than in the traditional low interactivity abstract version of the water jar task. In this ecosystem, problem solving traces a spatio-temporal trajectory evidenced by the changes in the agent-environment configuration. The cognitive psychology of problem solving must characterize the system, the agent embedded in it and the transformative coupling that uniquely enacts both.

Insight Problem Solving. Transformation or analytic problems have not exercised cognitive psychologists as much as so-called 'insight' or 'non-routine' (Mayer, 1995) problems. The solution to an insight problem poses an interesting empirical and theoretical challenge because the solution appears *not* to reflect the gradual transformation of proto solutions into a fully-fledged answer, but rather emerges with a certain velocity and clarity after much effort to overcome an impasse. Theoretical efforts from a number of different perspectives (Segal, 2004; Ohlsson, 1992; Fleck & Weisberg, 2013) converge on the importance of the initial problem representation and how that representation is restructured to offer the solution to the problem. Thus insight problems—as created under laboratory conditions—are designed to encourage a misleading initial interpretation of the problem, which must be abandoned and restructured into one that prompts the right operators to yield the solution (see Fig. 2 for a classic example).

Insert Figure 2 about here

The cognitive psychologist's deep commitment to intracranialism could not be clearer: "*Representation* occurs when a problem solver builds an internal mental representation of a problem that suggests a plan or a solution; *solution* occurs when a problem solver carries out a solution plan" (Mayer, 1995, p. 4, emphasis in the original). In other words, the world is represented; internal computational processes transform this representation; each representational state cues operators that further transform the representation until it spawns a mental plan of action, which then governs problem solving behavior. The psychology of reasoning thus aims to characterise a pure computational process, classically decoupled from a physical environment that instantiates it. A series of figures in Ohlsson's (2011) recent book *Deep learning* (Figs. 4.2-4.4 pp. 96-110) illustrates the conjectured processes by which activation among nodes that represent perceptual or semantic features of a problem is re-distributed as a result of negative feedback, that is when an impasse is experienced. The outside world is a source of information to be sure, but in classical information-processing terms, the science begins when the world is represented. An important methodological—and ultimately theoretical—consequence of this commitment to such a strong thinking-is-in-the-head position is an indifference, or partial blindness, to interactive sense-making activities, to problem representations that are distributed and that emerge from action- perception loops, which, in turn, shape and re-shape the environment to support problem solving behavior.

This is particularly clearly illustrated in the methodology employed by Fleck and Weisberg (2013) in their recent paper that offers a general theoretical framework of problem solving. The authors selected five insight-like problems and trained participants to provide verbal protocols as they laboured to find a solution. Remarkably, two of the problems were text-based problems, and three were presented with manipulable artefacts. The fact that some problems fostered interactivity and the dynamic reshaping of the problem presentation failed to exercise the authors at all (a similar methodological indifference is found in the series of problems selected in the Gilhooly and Fioratou's [2009] study on the role of working memory in problem solving). However, there is much evidence in Fleck and Weisberg's paper that restructuring is much more likely in high interactivity contexts. For example, from the information in their Table 2 (p. 450), across successful and unsuccessful problem solvers, 67% of the participants restructured their representation when working on the problems with artefacts, but only 18% did so for the problems without the artefacts. If one only looks at the successful participants, the presence of artefacts lead to a greater number of solutions through restructuring. Yet, Fleck and Weisberg (2013) never comment on these patterns in their data, and never address the main driver of restructuring, namely interactivity (see Vallée-Tourangeau, under review, for a more detailed analysis of the theoretical framework and the data reported in Fleck & Weisberg, 2013).

A distributed problem representation is structured and restructured along a contingent temporal and spatial path. An emergent distributed representation reflects the dynamic transactional coupling of an agent's mental space (cf. Malafouris, 2013, p. 101) and the shape of the physical environment in which she is embedded. As an example, imagine she is given 7 letter tiles—e.g., L, N, A, O, I, T, E (Maglio et al., 1999)—and asked to generate a 7-letter world. She could be an anagram maven and spurt out a 7-letter specimen through sheer mental zip. Most likely, however, she will start moving the tiles, re-arranging their order, experimenting with different stems and morphemes. The production of a new word lies in the intertwining of the agent's proto-hunches and the letter strings offered by the environment. Her actions need not be guided by a fully articulated plan, or indeed any plan at all (e.g., she could be playing with the tiles). She doesn't yet know what she is after: her representation of the candidate word is partial and incomplete, tethered to the letter string that she's just re-arranged. Her actions will exploit certain environmental affordances (e.g., in the form of certain letter strings, certain physical juxtapositions), themselves contingent on previous letter tile movements. She first generates T O N A L, and with a few additional movements she anticipates the possibility of T O E N A I L; the solution to her problem is thus distilled by interactivity. This simple illustration highlights an important methodological challenge: To understand insight, research should focus on the dynamic transformation of the agent's mental space in tandem with the transformation of the external environment. For this to happen, though, the research procedure employed must permit and encourage interactivity and a malleable problem presentation that can be transformed through an agent's action.

Interactivity. A well-known productive thinking task, initially developed by Knoblich, Ohlsson, Haider and Rhenius (1999; see also Van Stockum & DeCaro, 2013) involves matchstick arithmetic with Roman numerals. The participants are presented with a false expression such as 'II = II + II' and are invited to turn it into a true one by moving only one matchstick. In this particular example, the operator '+' can be transformed into the operator '-' and the additional 'stick' can be added to the numeral 2 to turn it into a 3 (viz. II = III – II). This is a productive thinking task because there is no well learned routine that can be applied to discover the solution. Knoblich et al. (1999) report the solution rates of different matchstick problems varying in difficulty. In this, and subsequent experiments (e.g., Knoblich, Ohlsson, & Raney, 2001) the experimental procedure makes sense only from a deeply entrenched commitment to an internalist-classical-computational view of cognition: The problems don't actually involve matchsticks as such; the arithmetic expressions are presented on a computer screen and participants announce their proposed solution to an experimenter. In their subsequent eye-tracking study, Knoblich et al. (2001) had participants place a 'bite bar' in their mouth to "increase the precision of the measurement by stabilizing the participants' head" (p. 1003). Thus, participants stared at an unmodifiable display. Whatever mentally simulated transformation, the perceptual feedback remains the same and considerable attentional resources must be deployed to maintain the simulated transformation in mind despite the conflicting, unchanging, sensory experience. Thus the information in an unchanging display is much different from the information that emanates from the pliant and dynamic external world. Whatever may be said in favour of the window that such an experimental procedure opens onto problem solving activity as they naturally occur in the world outside the laboratory, the feature that undermines their ecological plausibility is the absence of flexibility. The world outside is plastic: it is kneadable, malleable, modifiable. Reasoners naturally transform the environment when thinking, or rather transforming the environment is thinking: Interactivity matters. Imagine the same task but with moveable matchsticks and participants free to move the matchsticks when 'thinking'. Moving a matchstick, whether on the basis of a plan of action or action cued by perceptual features of the environment, transforms the arithmetic expression, hence the perceptual feedback, and may alter the epistemic potential of the environment. Thus the proto-hypotheses that coalesce in the reader's mind and the physical arithmetic expression are co-determined through action; the conscious explicit expression of a solution to the experimenter in this instance is a second order narrative primarily driven by first order action-perception loops. Moving the horizontal stick that composes the plus operator in the example above transforms the perceptual information as well as the reasoner's proto-hunches; the dynamic perceptual information transforms the reasoner's mental space, and both cue different actions that bind this contingent transactional co-existence.

Weller, Villejoubert and Vallée-Tourangeau (2011) designed an interactive version of the matchstick arithmetic task employed in Knoblich et al.: Participants were much more likely to solve the insight problems when given the opportunity to manipulate the matchsticks than when they looked at a static display of the problems. The behavior of participants in the low interactivity group was also interesting: they would touch the printed numerals or move their fingers across the printed arithmetic expression. Kirsh (1995) would call these movements complementary actions that helped participants simulate the matchstick movement in their mental space. In the interactive condition, however, participants could materialise these projections. A materialised projection can be manipulated and anchors new projections. Kirsh (2013) calls this the most fundamental feature of human cognition: “When we interact with our environment for epistemic reasons, we often interact to create scaffolds for thought, thought supports we can lean on. But we also create external elements that can actually serve as vehicles for thoughts. We use them as things to think with.” (p. 178). As the environment is modified, the perceptual feedback and action affordances are modified, cueing a different set of behavioural opportunities that further determine the spatio-temporal development of the agent-environment configuration. Engineering insight problem solving task where participants can modify the physical features of the problem presentation permits the qualitative and quantitative characterisation of such systemic dynamic configurations.

Cognitive Ecosystems

Representational effects are well understood and long appreciated by cognitive psychologists. To adapt the 'number scrabble' game described in Simon's (1996) *The sciences of the artificial*, imagine the Game of 15, where players take turn selecting numbers from 1 to 9; they do so without the support of pen or paper and with the goal of being first to string three numbers that sum to 15. In the absence of any artifacts and the possibility of projecting moves onto a grid of sorts (cf. Kirsh, 2009), the game quickly takes a toll on working memory. After a few turns, it is difficult for players to keep track of their own selections, all possible interim sums, not to mention the numbers chosen by their opponent: As a result the strategic number selection to block an opponent's progress or guarantee a win becomes very difficult. However, if one can conjure up in working memory a magic square (as illustrated in Fig. 3), where the sum of each row, column, and diagonal is 15, then the selection of numbers and the appreciation of which moves can draw or win the game is 'just' like playing noughts and crosses (e.g., if player A has already selected '4' and '5' then B must select '6'). Of course, the game of 15, for most people, is much harder to play than noughts and crosses, yet at one level of analysis, the game of 15 and noughts and crosses are isomorphic. This isomorphism encourages and reassures a strong internalist and computational explanation of behaviour. The bet is that the modeling efforts must be based on the deep abstract logical structure of the game. In other words, the fact that noughts and crosses is easier to play than the game of 15 is a secondary detail of implementation. However, the cognitive ecological realities can be very different across different instantiations of 'isomorphic' tasks. For example, across the two versions of the water jar volume measurement tasks (as discussed in the preceding section), the perceptual information, the tools and artefacts, action affordances and proprioceptive feedback configure different cognitive ecosystems, which engage and constrain cognitive agents in very different ways.

Insert Figure 3 about here

Zhang and Norman (1994) offered an elegant demonstration of representational effects in an analytic problem-solving task. On the basis of a formal analysis of the manner with which key features of a problem can be physically realised in the artefacts employed to present the problem, they predicted how certain Tower of Hanoi problem isomorphs might be quicker and easier to complete. The measures and analyses of problem solving performance presented in Zhang and Norman (1994) however ignore a fundamental aspect of thinking, namely that it reflects the evolution of an agent-environment system configured through interactivity. Zhang and Norman measure the time and the number of moves needed to solve the various Tower of Hanoi isomorphs as a function of which rule(s) or dimension(s) were externally represented, but the actual trajectory enacted with these different isomorphs was never mentioned, measured, or analysed. In addition, by using an analytic problem with a very clear goal state and a small set of operators, participants knew what the solution looked like: they were not looking for a solution, they were looking for means to achieve it quickly with the fewest moves. In contrast, insight problems have goal states that initially appear irreconcilable with the given information and the operators needed to transform progressive approximations of the goal state are unknown. We have argued that insight is distilled through interactivity, and it's important for researchers to develop a research methodology to capture the contingent trajectory wrought by interactivity.

The work reviewed in this paper stresses the importance of interactivity in enhancing reasoning and problem solving. However, the material presentation of a problem, the nature of the external resources recruited in the process of solving a problem, may also steer the thinking trajectory along unproductive paths; we have also observed this phenomenon in our work (Vallée-Tourangeau, Steffensen, Vallée-Tourangeau, & Makri, in preparation; Vallée-Tourangeau, under review). Complex systems can be unpredictable, and are susceptible to initial conditions. Unanticipated action affordances in the design of the artefacts in a problem solving experiment may encourage the reasoner to engage in unproductive behavior. We would argue, however, that these unproductive trajectories are also better understood from a systemic perspective that help researchers identify more precisely the nature of the external resources that best supports sound thinking.

Different problem isomorphs instantiated with different sets of artefacts that afford certain action opportunities specify a certain cognitive ecosystem. Implicit, perhaps, to much of our discussion thus far, is the question of which of these different cognitive ecologies is the canonical one that offers the best window onto a particular feature of human reasoning. However, such a question assumes that there is a pure or true reasoning ability that can be unlocked with the right experimental key. It is probably wiser to assume that a constellation of lab-based cognitive ecologies can be loosely grouped together on the basis of how well they approximate cognitive ecologies outside the laboratory. In addition, a commitment to a systemic perspective on reasoning undermines efforts to localize the proximate cause of reasoning performance in the cognitive agents themselves. The transactional forces that shape the different components of the cognitive ecosystem means that the system is ‘non-decomposable’ (Baber, Parekh, & Cengiz, 2014): No components of the cognitive ecosystem can be isolated and implicated independent of the other components that together configure the ecosystem (Steffensen, in press; Turvey & Shaw, 1999).

In reviewing inductive reasoning and problem solving (of the reproductive and productive kind), we’ve sought to encourage researchers to examine carefully the agent-context coupling that defines the cognitive ecosystem in which a certain reasoning task is completed. This systemic perspective offers a much richer source of insights about human reasoning, and pays off handsomely in empirical, methodological and theoretical terms. Empirically, the perspective offers inexhaustible heuristic value in pushing researchers to tinker with the parameters of the agent-environment coupling and in the process, unveil a broader spectrum of problem solving behaviours and reasoning performance. Dispositional conjectures and commitments to individualism promote a cognitive elitism that constrains the ability to envisage and engineer transformative agent-context interfaces. Methodologically a systemic perspective naturally encourages critical reflections on the laboratory procedure designed to offer a window on some features of human cognition. Research paradigms fossilize through feasibility and replicability pressures and then through the narrow parametric tweaking of proposed explanatory models. Eventually, the original methodological decisions made by early researchers are black-boxed and unchallenged. Wason’s (1960) original procedure has been employed repeatedly, each time demonstrating biased hypothesis testing behaviours that limit people's ability to discover the rule. The replication of the phenomenon from such a narrow set of methodological parameters reflect the type of biased hypothesis testing that the task was designed to illustrate in the first instance.

Theoretically, a systemic perspective invites researchers to rethink the nature of cognition, of representations, and the mechanisms that assemble a cognitive product (cf. Giere, 2006). Grand ontological pronouncements oblivious of the nature of the tasks faced by reasoners, the social and cultural dimensions of the system, and the different time scales over which cognition unfolds simply cannot accommodate the range and the dynamic ontogeny of the cognitive phenomena investigated by psychologists and social scientists. The cognitive products reviewed here—a problem representation, a hypothesis that describes sequences of numbers, or a solution to a problem—can be better described in terms of a spatio-temporal trajectory enacted through agent-environment transactional couplings (Vallée-Tourangeau & Vallée-Tourangeau, 2014). Along this trajectory, complexity and representational content snowball, and cognition evolves into a second order narrative of a sequence of first order action-perception loops. Ontological considerations must thus reflect ontogenic and scalar constraints. These theoretical considerations suggest that a productive way to proceed in the psychology of reasoning is to adopt a more qualitative description and analysis of participants' behavior over time as they work on the problem (e.g., Steffensen, 2013). Such a shift from individual methodologism to systemic methodologism would leave researchers in a better position to document the mutually transformative effect of interactivity as participants reshape the physical configuration of the problem constituents which in turn shape the distributed representation of the problem.

In psychology experiments, researchers manipulate variables to study the relationships between specific events and their consequences. In cognitive psychology, a typical experiment involves the presentation of different verbal, iconic or auditory stimuli and measuring the effect of a change in stimulus on a behavioural response such as a reaction time or a performance level (cf. Järvilehto, 1998). The tool of choice for running experiments in many laboratories is a personal computer as they allow for precise control over the presentation and timing of stimuli and the responses measured. Another common tool is simply a paper-based questionnaire presenting the stimuli and asking participants to respond using a pencil. The apparatus used in these experiments, however, is likely to bias our understanding of what cognition is and how it works. The implicit limitation of both the computer-based or paper-based stimuli presentations is that they severely limit participants' opportunities to handle the information presented to them. This may be seen as an asset for the experimentalist who has adopted an internalist perspective on cognition and, consequently, whose primary concern is to control for extraneous influences. From a systemicist's perspective, however, this choice of apparatus offers but a limited cognitive ecosystem which imposes artificial limitations on the affordances available to the agent for performing the cognitive task at hand and, consequently, may offer a biased window onto the feats a reasoning agent can achieve and how she may achieve them. The alternative is to consider the affordances offered by the tasks and consider using stimuli that can be manipulated or handled by the participants while they think, either by using actual artefacts in lieu of static verbal descriptions or interactive diagrams in lieu of static representations and interactive conversations in lieu of controlled auditory stimuli. For example, in Vallée-Tourangeau and Payton (2008), the interface afforded the active manipulation of the diagrammatic representation of the triplets generated. This, in turn, enabled participants to be actively engaged in a cycle of cognitive transactions between what they thought in their mind, what they did in the world, and what the world offered back. We believe that untethering participants and offering them the opportunity to think through their hands may have a transformative effect on our understanding of cognition. As the systemicist's approach matures, this will allow researchers to gain a finer understanding of which types of affordances have the potential to transform cognition and how they may do so.

References

Baber, C., Parekh, M., Cengiz, T. G. (2014). Tool use as distributed cognition: Howe tools help, hinder and define manual skill. *Frontiers in Psychology, 5*, Article 116, 1-16

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition, 108*, 652–661.

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.

Dominowski, R. L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg and J. E. Davidson (Eds.), *The nature of insight* (pp. 33-62). Cambridge MA: The MIT Press.

Evans, J. St. B. T. (in press). Reasoning, biases and dual processes: The lasting impact of Wason (1960). *Quarterly Journal of Experimental Psychology.*

Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology*, 25, 436-463.

Gale, M., & Ball, L. J. (2006). Dual-goal facilitation in Wason's 2-4-6 task: What mediates successful rule discovery? *Quarterly Journal of Experimental Psychology, 59*, 873-885.

Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking and Reasoning*, *15,* 355-376.

Giere, R. (2006). *Scientific perspectivism*. Chicago: University of Chicago Press.

Järvilehto, T. (1998). The theory of the organism-environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science, 33,* 321-334.

Jensen, J. (1960). On the Einstellung effect in problem solving. *Scandinavian Journal of Psychology, 1*, 163–168.

Kirsh, D. (1995). Complementary strategies: Why we use our hands when we think. In J. M. Moore & J. L. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 212-217). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Kirsh, D. (2009). Projection, problem space and anchoring. In N. Taatgen, H. van Rijn, J. Nerbonne & L. Schomaker (Eds.), *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society*. (pp. 2310-2315). Austin, TX: Cognitive Science Society.

Kirsh, D. (2013). Thinking with external representation. In S. J. Cowley, & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 171-194). London: Springer-Verlag.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*, 1000-1009.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, *25*, 1534-1555.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monograph, 54* (Whole No. 248).

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, NJ: Lawrence Erlbaum Associates, Publishers.

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. London: The MIT Press.

Mayer, R. E. (1995). The search for insight: Grappling with Gestalt psychology's unanswered questions. Insight and problem solving. In R. J. Sternberg and J. E. Davidson (Eds.), *The nature of insight* (pp. 3-32). Cambridge MA: The MIT Press.

Ohlsson, S. (1984). Restructuring revisited II: An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology, 25,* 117-129.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (pp. 1- 44). Hemel Hempstead, UK: Harvester Wheatsheaf.

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

Popper, K. (1935/1992). *The logic of scientific discovery*. London: Routledge.

Ross, L .D., Amabile, T. A., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology, 35*, 485-494.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, *16*, 141-148.

Steffensen, S. V. (2013). Human interactivity: Problem-solving, solution probing and verbal patterns in the wild. In S. J. Cowley, & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 195-221). London: Springer-Verlag.

Steffensen, S. V. (in press). Distributed language and dialogism: Notes on non-locality, sense-making and interactivity. *Language Sciences.*

Simon, H. A. (1996). *The sciences of the artificial* (3rd edition). Cambridge, MA: The MIT Press.

Turvey, M. T., & Shaw, R. E. (1999). Ecological foundations of cognition I: Symmetry and specificity of animal-environment systems. *Journal of Consciousness Studies, 6* (11-12), 95-110.

Vallée-Tourangeau, F. (under review). *Insight, interactivity and materiality.*

Vallée-Tourangeau, F., Austin, N. G., & Rankin, S. (1995). Inducing a rule in Wason's 2-4-6 task: A test of the information-quantity and goal-complementarity hypotheses. *Quarterly Journal of Experimental Psychology*, *48A*, 895-914.

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

Vallée-Tourangeau, F., & Payton, T. (2008). Graphical representation fosters discovery in the 2-4-6 task. *Quarterly Journal of Experimental Psychology, 61*, 625-640.

Vallée-Tourangeau, F., Steffensen, S. V., Vallée-Tourangeau, G, & Makri, A. (in preparation). *Insight in cognitive ecosystems*.

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

Vanderhenst, J.-B., Rossi, S., & Schroyens, W. (2002). When participants are not misled they are not so bad after all: A pragmatic analysis of a rule discovery task. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society* (pp. 902-908). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Van Stockum, C., & DeCaro, M. S. (2013). The path less taken: When working memory capacity constrains insight. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the Thirty-Fifth Annual Conference of the Cognitive Science Society* (pp. 3633-3638). Austin, TX: Cognitive Science Society.

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

Weller, A., Villejoubert, G., Vallée-Tourangeau, F. (2011). Interactive insight problem solving. *Thinking & Reasoning*, *17*, 429-439.

Wertheimer, M. (1959). *Productive thinking*. New York: Harper & Brothers Publishers.

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

Figure Captions

Figure 1. Three types of triples in the 2-4-6 task represented graphically as in Vallée-Tourangea and Payton (2008): (a) ‘6-8-10’; (b) ‘6-12-10’; (c) ‘1-5-13’.

Figure 2. A Max Wertheimer problem: Find the surface area of the figure composed of a parallelogram on top of a square. The initial problem representation (a) creates an impasse for most participants (see Segal, 2004). A restructuring of the representation in terms of overlapping triangles (b) and then into a rectangle (d) creates a much simpler problem that is more easily solved (Ohlsson, 1984).

Figure 3. The game of 15: Winner is the first to pick three numbers that sum to 15. If player A's selected '4' and '5' then B must select '6'. The game is isomorphic with noughts and crosses.

Figure 1

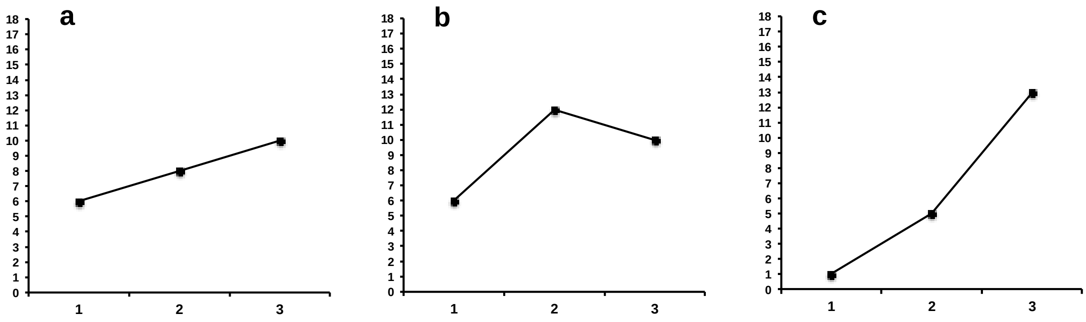

Figure 2

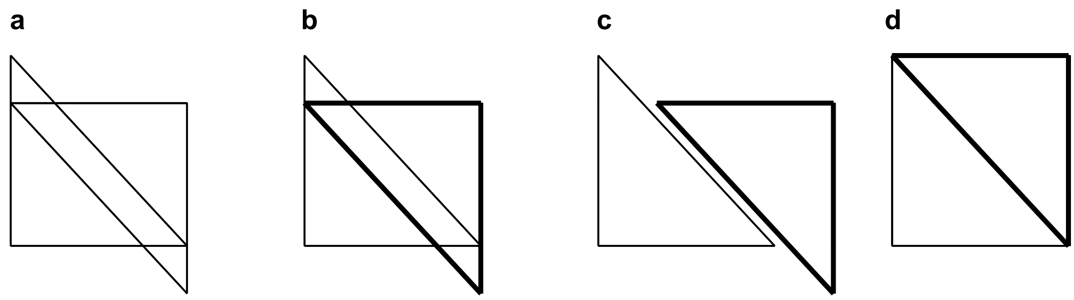

Figure 3

4	3	8
9	**5**	1
2	7	6