Conceptual Transformation and Cognitive Processes in Origami Paper Folding

Tenbrink, T.; Taylor, H.A.

Journal of Problem Solving

DOI:
10.7771/1932-6246.1154

Published: 07/07/2015

Journal of Problem Solving

Conceptual Transformation and Cognitive Processes in Origami Paper Folding

Thora Tenbrink[1] and Holly A. Taylor[2]

[1]Bangor University, [2]Tufts University

Correspondence:
Correspondence concerning this article should be addressed to Thora Tenbrink, Bangor University, Bangor, Gwynedd, LL57 2DG, UK, or via email to t.tenbrink@bangor.ac.uk.

Acknowledgments:
This research was generously supported by the DFG (SFB/TR 8 Spatial Cognition, project I6-[NavTalk]), the Hanse Institute for Advanced Studies, and the SILC Spatial Intelligence and Learning Center. We are also grateful for the support of our very reliable student assistants.

Keywords:
problem solving, instructions, text interpretation, cognitive processes, verbal data analysis, reconceptualization

Research on problem solving typically does not address tasks that involve following detailed and/or illustrated step-by-step instructions. Such tasks are not seen as cognitively challenging problems to be solved. In this paper, we challenge this assumption by analyzing verbal protocols collected during an Origami folding task. Participants verbalised thoughts well beyond reading or reformulating task instructions, or commenting on actions. In particular, they compared the task status to pictures in the instruction, evaluated the progress so far, referred to previous experience, expressed problems and confusions, and—crucially—added complex thoughts and ideas about the current instructional step. The last two categories highlight the fact that participants conceptualised this spatial task as a problem to be solved, and used creativity to achieve this aim. Procedurally, the verbalisations reflect a typical order of steps: *reading—reformulating—reconceptualising—evaluating*. During reconceptualisation, the creative range of spatial concepts represented in language highlights the complex mental operations involved when transferring the two-dimensional representation into the real world. We discuss the implications of our findings in terms of problem solving as a multilayered process involving diverse types of cognitive effort, consider parallels to known conceptual challenges involved in interpreting spatial descriptions, and reflect on the benefit of reconceptualisation for cognitive processes.

Origami is the well-known Japanese art of creating 3-D objects by folding paper in a particular manner and order. Often, this is achieved by following written instructions supported by pictures, for example, from a book or webpage. How do people interpret abstract action descriptions to create a concrete object resembling what is shown in a picture? Anyone who has ever struggled with the challenge of folding Origami, or used any kind of manual to assemble an object or comprehend a newly acquired technical device, will be familiar with potential misinterpretations and conceptual traps. Learning a new procedure based on pictures and text may represent a problem requiring considerable mental effort to solve.

Some cognitive complexity arises when conceptually transfering from an abstract medium toward concrete actions. Moreover, language and depictions, even together, as communication media are notoriously underspecified, leaving more room for interpretation than one might desire (Carston, 2002; Hegarty & Just, 1993; Van Deemter & Peters, 1996). In general, if intended actions need instructions, then there is a problem to solve, and instructions can support the task. Even with instructions, subtle decisions and individual conceptualisations engaged during problem solving mean that the outcome may not always be successful.

Research in problem solving in general has mainly focused on identifying creative problem solving, for instance, in order to propose adequate sets of step-by-step instructions (e.g., Anderson, Douglass, & Qin, 2004). However, the act of following instructions has not received extensive research attention. Since instructions guide people along a conceptual path, the need for creativity and/or individual strategies might seem limited.

In this paper, we challenge this assumption by treating a complex instruction-based task, namely Origami folding, as a problem needing a solution via a range of conceptual steps. We start by reviewing the role of operations and cognitive strategies in the problem solving literature, and then consider insights from research examining text and picture comprehension, particularly in the spatial domain. Then we report our study in which participants folded an Origami object (a flower stem) while thinking aloud. Our analysis first addresses the extent to which participants' verbalisations reflect creative problem solving processes beyond reading or reformulating and expressing task execution, and then focuses on the types of conceptual steps represented in the verbalisations. We highlight how participants iteratively interpret and reconceptualise each folding step until satisfied with the produced object. Then we focus on the reconceptualisation process as a main component of the complex problem solving of Origami.

http://dx.doi.org/10.7771/1932-6246.1154

MENTAL PROCESSES IN PROBLEM SOLVING TASKS

Following the seminal approach by Newell and Simon (1972), human problem solving means conceptually breaking down a problem into separate and manageable steps or operations. In his representative account of the state of the art, Anderson (2004, p. 245) characterises problem solving as "goal-directed behavior that often involves setting subgoals to enable the application of operators." Here, "the term operator refers to an action that will transform the problem state into another problem state. The solution of the overall problem is a sequence of these known operators," and "the challenge is to find some possible sequence of operators in the problem space that leads from the start state to the goal state" (Anderson (2004, p. 245). Accordingly, much of the problem solving literature addresses how people identify problems and operators to solve them, and how these operators are ordered into a sequence of actions so as to reach a suitable solution, mediated by expertise (Chi, Glaser, & Rees, 1982). This is reflected in the relevant literature such as Newell and Simon (1972), and more recently in many contributions in the *Journal of Problem Solving*—compare discussions in Carruthers and Stege (2013) and Fischer, Greiff, and Funke (2012), and in introductory reviews such as Anderson (2004), which focus on the complex high-level operations that need to be mentally organised, based on the range of possible actions and problem states.

It is in this area that think-aloud protocols as data sources have been most successful (Ericsson & Simon, 1984). This is because the identification and ordering of operators happens on a high cognitive level; solution steps to a complex problem are verbalisable to a great extent, as they are consciously accessible and can be adequately represented in language. A vast amount of problem solving research drawing on verbalisation data confirms this (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Gero & McNeill, 1998; Kuipers, Moskowitz, & Kassirer, 1988; Ritter & Larkin, 1994; Van Gog et al., 2005), in spite of issues about the verbalisability of specific kinds of problems such as those involving reading (Afflerbach & Johnston, 1984) or instantaneous insights (Schooler, Ohlsson, & Brooks, 1993).

One way of representing a solution path is by way of a process model (Fischer, Greiff, & Funke, 2012; Myers, Gluck, Gunzelmann, & Krusmark, 2010), for instance, using a cognitive architecture such as ACT-R (e.g., Gugerty & Rodes, 2007). Such models can be used as a basis for producing efficient and cognitively supportive instructions (Anderson et al., 2004). The emphasis on creating supportive instructions strongly suggests that not all instructions can be followed in a cognitively straightforward manner. As such, our focus in the present paper lies in the opposite direction—understanding how humans deal with existing instructions for a complex task. This differs from the kinds of problems addressed in problem solving research typically (or perhaps always), that is, ones for which the problem solvers do not have access to instructions or manuals.

Instructions reduce a given problem considerably, by offering a breakdown of the original problem into separate solution steps (operations) delineating a predetermined solution path. What remains is a more fine-grained cognitive challenge of moving "from a declarative representation and a slow interpretation of the task to a smooth, rapid procedural execution of the task" (Anderson et al., 2004, p. 1046). Highlighting this challenge, Anderson, Kline, and Lewis (1977) as well as Ball (2004) proposed cognitive models for language processing in general. However, to our knowledge, the interpretation of instructions (in terms of guided actions) has not been addressed directly as a problem-solving task for which cognitive processes can be modeled.

A possible reason for this is that cognitive processes in following instructions are not expected to be accessible via existing measures such as behavioural performance outcomes or verbalisation protocols. Transforming a given declarative representation into action may conceivably involve entirely low-level cognitive processes, since no further identification of problem solving steps is required. If that is the case, humans who follow instructions to solve a problem should not have much to verbalise beyond reading and perhaps reformulating the instructions. However, anyone who has tried to follow complex instructions would likely attest that problem solving opportunities arise in this context.

In this paper we address this assumption, and ask if following instructions for a complex and cognitively challenging task such as Origami folding may elicit thought processes that can, to some extent, emerge in task-concurrent verbalisations. In order to see what kinds of challenges may be involved in following instructions, we now turn to research on the interpretation of textual and visual representations.

INTERPRETATION OF TEXT, PICTURES, AND INSTRUCTIONS

Reading a text activates a number of mental processes toward comprehension. According to a bottom-up model of discourse comprehension proposed by Kintsch (1988), spreading activation of concepts based on linguistic cues leads to the construction of a mental representation of the text. Zwaan and Radvansky (1998) further suggested that readers construct a coherent situation model that integrates every newly read clause with the information accumulated so far. This process involves complex interactions of long-term memory retrieval and short-term memory activation. Furthermore, intricate grounding processes with respect to temporal and spatial domains are necessary for the situation model to be consistent. Readers may develop a mental image

(Kosslyn, 1980) of the textual content, which may amount to a simulation of the situation (Barsalou, 1999), representing details such as spatial structures in the visual field (Bergen, Lindsay, Matlock, & Narayanan, 2007).

Common to all discourse comprehension approaches is the insight that the original text formulation serves as a trigger for accessing conceptual frames, logical and commonsense based inferences, and knowledge elaborations that are not directly expressed in the text. Readers quickly identify a message's gist and typically cannot remember the original wording after a very short time (Bransford & Franks, 1972; Sachs, 1967). While readers quickly derive a suitable context-dependent interpretation from their mental model of the situation related through a text, more complex inferences require further cognitive effort and are not as readily incorporated (Garrod, 1985). This effect is similar to problem solving in general in that intuitive and effortless reasoning is replaced by meta-cognition and higher-level conscious processes (only) when particular challenges or problems occur (Alter, Oppenheimer, Epley, & Eyre, 2007).

Pictorial information can support text interpretation. For instance, Bransford and Johnson (1972) found that people better recalled a text illustrated by a picture that provided essential context. In terms of comprehension, picture and diagram interpretation proceeds similarly to reading comprehension in that the gist and conceptual frame are identified quickly, guiding attention towards relevant aspects to serve a particular purpose (Franconeri et al., 2012; Henderson, 2003). The process is facilitated by the fact that depictions can resemble the mental abstractions necessary for remembering and visualising relationships (Tversky, 2011). When combining pictures and text, comprehension can be hampered or supported by particular features of spatial integration, visual complexity, relevance, and the like (Florax & Ploetzner, 2010). Altogether, the comprehension of descriptions and depictions draws on similar but not identical principles, which in the ideal case work together to allow for a thorough understanding (Schnotz, 2002; Schnotz, Baadte, Johnson, & Mengelkamp, 2012).

Since different contexts and contents call for different representations, identifying an ideal form remains a challenge in every case. Ultimately, no representation is complete or directly accessible to the human mind; intricate comprehension processes are required to gain an adequate interpretation. Different modes of representation affect the distribution of cognitive load in systematic ways, depending on the represented content and its adequacy relative to the recipient's level of expertise (Cook, 2006). For instance, the extent to which complete and detailed information is necessary or beneficial for a reader depends on their background. With a high level of previous knowledge to draw on, readers benefit from the challenge posed by less complete representations that call for deeper processing. Texts that leave room for the readers' inferences support a more thorough understanding due to the increased activation of interpretive processes and linking to one's knowledge base (McNamara, Kintsch, Butler Songer, & Kintsch, 1996). Relatedly, different types of learning materials are useful for different purposes (Belenky & Schalk, 2014); while initial learning is enhanced by grounding in background information, transfer is easier when abstracting across contexts (Gick & Holyoak, 1980). However, learners differ in the extent to which they can generalise from examples. Crucially for our purposes, learners who successfully generalised provided explanations for themselves while reading, displaying their deep understanding, more than those who failed to generalise (Chi et al., 1989).

Comprehending instructions and manuals involves these general interpretation processes (Franck, 2004) with their complex interplay of context, represented information, background knowledge, and expertise, plus the challenges of resolving references to relevant objects (Weiß, Pfeiffer, Eikmeyer, & Rickheit, 2006), and transforming the information towards a practical purpose—actions to be undertaken in the real world (Daniel & Tversky, 2012). Paralleling the more general findings on text comprehension, Marcus, Cooper, and Sweller (1996) argue that the addition of diagrams can reduce cognitive load, making instructions easier to follow. Mediated by their ability and expertise in the subject area, readers construct a mental model by incrementally combining local with global information (Hegarty & Just, 1993). This is supported by situation-based affordances provided through experiential (non-propositional, non-abstract) background knowledge (Glenberg & Robertson, 1999). Real-world objects and displays offer visual feedback cues supporting action directly, reducing memory load and instantaneously suggesting possible actions (Larkin, 1989).

The processes and requirements involved with following instructions have been quite thoroughly researched in the context of route descriptions. For instance, Lovelace, Hegarty, and Montello (1999) proposed elements that make up "good" route directions. Completeness, mention of segments and turns, and particular types of landmarks contributed to route description quality ratings. Additionally, Allen (2000) showed that preserving the natural order and focusing on action information at choice points is important, as is taking the addressee's knowledge into account (this also affects the route planned, cf. Hölscher, Tenbrink, & Wiener, 2011). While visual information such as maps is just as useful for wayfinding as verbal route descriptions (Meilinger & Knauff, 2008), Lee and Tversky (2005) suggest that adding visual landmark information supports comprehension, in line with the insight that visual imagery can promote reasoning, especially in spatial settings (Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002; Tversky, 2011).

Particular challenges arise where spatial descriptions are underspecified or ambiguous, as is frequently found. The analysis of *dialogues* provides hints about the mental activity engaged in such cases. For instance, Tenbrink, Coventry, and Andonova (2008) found that addressees frequently suggested reformulations of, or additions to, spatial descriptions. Such reconceptualisations arise because of complex inference processes involved in spatial settings, as specified by Krause, Reyle, and Schiehlen (2001). Muller and Prévot (2009) identified types of feedback addressee's provided as a function of the information given by the speaker, enabling the dialogue partners to negotiate spatial representation challenges. Overall, the dialogic patterns reflect the need to integrate spatial descriptions into a coherent spatial mental model. Together, these results point to a high amount of creativity and cognitive processing on several levels (direct and effortless, as well as mediated and meta-cognitive) when following verbal descriptions of space. In other words, they point to the need for problem solving when following instructions.

FOLLOWING ORIGAMI INSTRUCTIONS: A PROBLEM-SOLVING TASK?

For tasks like Origami folding, few studies have explored mental processes involved in interpreting illustrated instructions. In face-to-face instruction of Origami, learners rely intensely on the instructor's gestures and actions to support the learning process (Furuyama, 2000). Because Origami can enhance spatial thought processes, training can lead to student gains when implemented into school curricula (Higginson & Colgan, 2001; Robichaux & Rodrigue, 2003; Taylor & Hutton, 2013). Algorithms for automatically interpreting graphical depictions of the folding process highlight the conceptual challenges and routines involved (e.g., Shimanuki, Kato, & Watanabe, 2003). While Sabbah (1985) provided a connectionist model for recognising line drawings of Origami objects, to our knowledge, the problem solving stages or conceptual steps of following Origami instructions have not been addressed.

Our aim in this paper is to provide insights about higher-level cognitive processes involved with interpreting illustrated instructions for folding a complex 3-D object. Rather than attempting to capture the finer processes involved in reading and picture comprehension, we focus on procedures and patterns reflected in think-aloud protocols, collected while following Origami instructions, and address patterns of variability in relation to individual and situation specific differences. Based on the research summarised above, we contrast alternative outcomes of our study:

- If instructions already spell out the main problem solving components typically found in verbalisable reports and thus accessible on a high level of cognition, and if there is no actual problem left to solve, then participants should simply follow instructions step by step, and carry out the task as outlined. Verbalisations would then consist of *reading* and slightly *reformulating* or adapting the instructions during the reading comprehension process, and *commenting* on how the given task is put into action.
- If following instructions is a problem to solve in itself, this should be expressed in the think-aloud protocols in terms of creative thinking or *additional ideas* that are not expressed in the instructions. Furthermore, participants might express *problems* in carrying out the task, and verbalise considerations as to how they might be solved.

Whether or not following Origami instructions can be seen as a problem solving task might differ depending on various factors. We expect variation based on participants' Origami *experience*, and we expect instruction steps to differ in terms of difficulty. These factors should be reflected in the verbal protocols, revealing how the conceptual challenge of following instructions is met according to the diverse factors involved, and what types or parts of instructions are particularly challenging.

Beyond identifying the existence of the relevant verbalisation types and indicators of the phenomena just outlined, we ask (qualitatively) how these thoughts are expressed in language, and what kinds of problem solving strategies and relevant verbalised concepts may occur, as the previous literature does not provide a sufficient basis for making direct predictions in this regard.

In response to the assumption that cognitive processes involved in interpreting instructions may be too low-level to be captured in think-aloud protocols, we employ *Cognitive Discourse Analysis* (CODA; Tenbrink, 2015) to address systematic features of the data. CODA was developed to capture deeper insights into cognitive processes, including those speakers might not be able to consciously verbalise, but nevertheless emerge in systematic patterns of verbalisations. The methodology extends the seminal approach to verbal protocol analysis by Ericsson and Simon (1984) by taking a closer look at the features of the language used to express thoughts and cognitive processes captured in a verbal protocol. The rationale behind this approach is that speakers make specific choices out of the more general network of lexicogrammatical options at their disposal. Such choices are meaningful in ways that speakers may not be aware of; for instance, they reflect a particular conceptual perspective and granularity level that appears natural to the speaker, but is in no way predetermined by the task. In this paper, the main CODA-based contribution concerns speakers' choices of spatial terms that were not directly part of the verbal instruction given to them, expressing their conceptual creativity while doing the task.

ORIGAMI STUDY

PROCEDURE

This study was reviewed and approved by the Tufts University Institutional Review Board. Twenty-four Tufts University undergraduates (14 male, 10 female), all native English speakers, participated in this study after having been fully informed about the procedures. They were trained to *think aloud* (see Appendix A), following suggestions by Ericsson and Simon (1984). Then their first task was to fold the Origami tulip, first the stem and then the blossom, following instructions provided on a computer screen. During these tasks they thought aloud while following the instructions (Appendix B). Participants could move through the instructions at their own pace, scrolling back and forth as they saw fit. The experimenter gave no advice except in the case of being stuck following a mistake. In such cases the experimenter provided a simple hint to reconsider the previous folding step. In cases of inactivity or silence, participants were encouraged to go on trying and to keep thinking aloud. Also, the experimenter provided positive feedback. A pilot test showed that, due to the considerable challenge of this task, such encouragement was vital. In spite of these adjustments, which were necessary to ensure a smooth task procedure and an actual outcome of each participant's efforts, it was made clear that there was to be no interaction about the task. The *think-aloud* expectation was transparent to the participants, who accordingly did not address the experimenter while verbalising their thoughts.

The participants' second task was to determine, in a series of trials, which of three Origami objects matched the crease patterns of an unfolded object. Finally, they completed three spatial abilities tests: the redrawn Vandenberg and Kuse mental rotation (Peters et al., 1995), mental paper folding (Shepard & Feng, 1972), and the Santa Barbara Solids Test (Cohen & Hegarty, 2012). We focus here on the cognitive processes reflected in verbalisations while folding the stem, without action coding (see Taylor & Tenbrink, 2013, for a different analysis of the same data set). The instruction for this task (represented in Appendix C) showed 13 folding steps as pictures associated with a brief textual instruction (e.g., "Put the paper in front of you, with the point toward the top").

RESULTS

Participants took between 3:05 and 10:07 minutes to fold the flower stem (mean = 04:54; standard deviation = 1:36). Eleven of the 24 participants received no hints by the experimenter, and the most hints given were four (mean = 1.16; standard deviation = 1.34). While folding, participants varied considerably with respect to how much they verbalised, producing between 113 and 1,738 words each (mean = 402.38; standard deviation = 337.75).

Folding success was assessed by independent ratings (7-point Likert scale) of the photographed stems. A separate group of 25 Tufts undergraduate students, who were not informed about the major goals of this study, rated each photograph. They rated success by comparing the photographs to the Origami instruction picture (see Appendix B), indicating the perceived similarity. Ratings ranged from 1.28 to 5.48 (mean = 3.97; standard deviation = 0.90). In other words, Origami folding results were judged as quite varied, covering almost the full range from failure to considerable success, although none of the resulting stems were unanimously considered entirely successful.

As would be expected, success ratings were marginally negatively correlated with the number of hints ($r = -0.37$, $n = 24$, $p = .073$). More interestingly, success was reliably negatively correlated with the number of words *read* ($r = -0.44$, $n = 24$, $p < .05$) rather than produced in more creative ways (see more detailed analysis of verbalisations below). That is, the more successful people were, the less they read instructions aloud. Apart from that, success was not related to any of our analysis criteria (including time to fold the stem), and will therefore not be further addressed as a determining factor in the analysis of the problem solving process as expressed in the verbalisations. Verbosity (i.e., the total number of words produced by a participant), for instance, was not related to folding success ($r = 0.081$, $n = 24$, $p>.05$), although it correlated with time to fold the stem ($r = .674$, $n = 24$, $p < .01$). No effects of gender emerged for any of our analysis criteria.

CONTENT CATEGORIES AND VERBAL CREATIVITY

All think-aloud data produced while folding the stem were transcribed and segmented into units containing a single thought or piece of information, such as "um, alright so I'm just trying to make sure it's as close to the fold as possible." Each unit was annotated in relation to the specific folding step (cf. Appendix C) to which it belonged.

As our first analysis goal, we explored the extent to which the verbalisations exhibited creative thought, as opposed to directly following the instructions. To assess this, we associated the content of units, or partial units if appropriate, with one of the following operationalised categories:

- *Reading task description:* parts that are read aloud or repeated from the written instruction about the relevant folding step.
- *Reformulating description without new thoughts:* conveying the same content as the instruction in a different syntactic or lexical form.
- *Additional ideas about a step:* introducing new ideas in describing this step. These were further subcategorised into the following (not mutually exclusive) types:

  - *orientation* of the paper to be folded
  - *object quality:* trying to get a nice result
  - *alignment* of the folds or edges with respect to each other
  - *crease quality:* making nice and sharp folds
  - *comparison to instruction:* trying to match participant's own result with the instruction (including the picture)
  - *within-step repetition*: doing the same action twice (e.g., for left and right sides) within a folding step
  - *across-step repetition*: the current step repeats a previous one (i.e., is described as identical)
  - *across-step difference:* the current step is compared to a previous one, identifying a difference
  - *spatial description:* patterns in the current status of the object
  - *adding semantics:* associating meaning with some part(s) of the current status of the object
  - *other.*
- *Evaluation:* the speaker evaluates their own work or progress so far in general terms, beyond the current folding step.
- *References to background knowledge:* for example, noting patterns based on experience.
- *Expression of problems:* considering how to do this step, expressions of matching problems, and so on.
- *Task communication:* the participant seeks confirmation about the procedure, comments on general aspects, refers to action (including looking at the pictures), explicitly starts the next step, or evaluates the instructions (as in "that makes sense").
- *Other:* anything ambiguous or not fitting into the previous categories.

As discussed above, we predicted that *Reading task description*, *Reformulating description without new thoughts*, and *Task communication* categories would reflect simple text interpretation and relevant action. All other categories go beyond this basic instruction-following process and were identified post hoc. They therefore represent a qualitative analysis of the types of thoughts participants verbalised.

Annotations were complete (all verbalisations were categorised) and (by our definition) mutually exclusive (i.e., no partial unit was associated with more than one main category, although the subcategories within the category *Additional ideas about a step* were not mutually exclusive). Annotation was achieved through an iterative multi-annotator coding process, ensuring optimised operationalisation of annotation definitions through repetitive in-depth scrutiny of the data, as well as consistency in coding by revisiting each data set multiple times as required. Following preliminary annotation by two independent student assistants, the process was only declared complete after both authors agreed with every instance of the annotations suggested by the students, following extended discussions of individual cases where needed. This iterative process was considered more adequate to the nature of this particular data set than a quantitative assessment of an inter-coder reliability measure (which is more typical).

Verbalisations coded as *Additional ideas* and *Expressions of problems* in particular reveal the conceptual issues associated with the Origami task (see Table 1 for examples of *Additional ideas*). Twelve of the 24 participants produced spatial descriptions such as "there's a straight line across here at the top," "making it more narrow," "touching the middle," "I have a triangle," "that one is horizontal," and so on, revealing that they identified spatial patterns within the folding process and resultant objects. This reflects a reconceptualisation of the original Origami instruction. Altogether there were 50 spatial descriptions of this kind.

Descriptions like these involved spatial vocabulary not included in a particular step's original instruction. To operationalise and verify this intuitive, content-based impression, we identified all instances of spatial terms used in relation to an instructional step, but not included in the relevant

Table 1.
Examples (taken from various individuals) for reconceptualisations categorised as *Additional ideas*

Instruction step no.	Instruction	Utterance	Subcategory
2	Fold the left corner over to the right one, and firmly straighten out the fold.	so that I get a triangle	Spatial description
5	Fold the bottom edges onto the midline.	this looks like a crane	Adding semantics
5	Fold the bottom edges onto the midline.	just take one corner and it's gonna go down a little bit	Spatial description
8	Fold the lower tip onto the upper one.	it looks like I have to match the height	alignment; comparison to instruction
11	Fold it back, and then diagonally to the left.	try and make it symmetrical	alignment

instruction (referred to as "new spatial terms" for short). We found a reliable correlation of "new spatial term use" with the "spatial description" subcategory within the *Additional ideas* category ($r = 0.57$, n = 24, $p < .01$; see also Taylor & Tenbrink, 2013). Thus, participants read the instruction and associated different spatial (and related) concepts with the action described, and expressed this in "new spatial terms." This provides insight into their cognitive flexibility in dealing with this task. Figure 1 shows the number of occurrence of the 25 most frequently used "new spatial terms," along with how many participants used each term. The most frequent term used in this way was *side*; 17 participants used it 104 times in situations where it was not part of the instruction. The remaining terms used, along with their frequency of occurrence, can be found in Appendix D. This impressively wide range of spatial terms highlights the creativity of thought employed by our participants.

As further illustrated in Figure 2, participants varied considerably in the extent to which they used "new spatial terms" throughout their verbalisations; counts ranged from 1 to 130 (mean = 21.38, median = 15). The production of spatial terms by individuals was (expectedly) correlated with the overall number of words produced ($r = 0.88$, n = 24, $p < .01$) as well as with other subcategories of *Additional ideas*)—clearly, the more verbose participants were, the more creative they became in their (spatial) language use. Also, use of new spatial terms was correlated with previous experience, as we report in more detail below.

Importantly, participants were not necessarily repetitive in their reformulations; each instruction step contained its own challenges and could therefore lead to new reconceptualisations and (as a consequence) different term use. To illustrate this, a closer look at the highest scoring dataset (130 new spatial terms) reveals that this participant produced 15 different spatial nouns: *angle, baselines, corner, crease, direction, end, edge, flap, line, position, shape, side, symmetry, three dimension, way.* In addition there were 7 different verbs: *bisect, end up, go, intersect, match up, switch, turn,* and 30 other spatial terms: *along, around, at, back, center, close, diagonal, down, even, flat, halfway, here, in, in half, into, lopsided, on, open, opposed, out, outside, over, overlap, straight, symmetrical, three dimensional, to, toward, up, vertical.* So, in total, this participant produced 52 different spatial terms, each about three times on average, to total 130. Of these terms, only 3

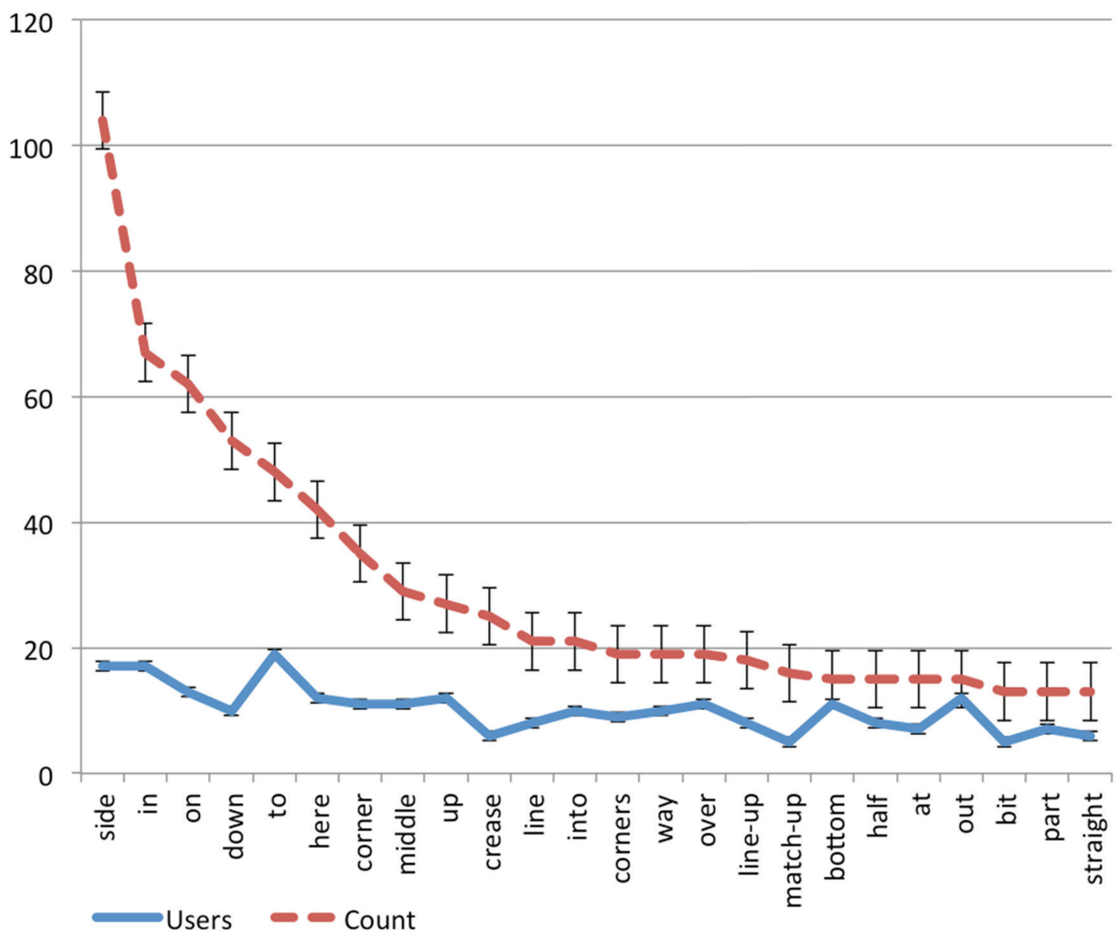


Figure 1. Frequency of "new spatial terms" used, along with the number of individuals who used the term at least once in a creative way.

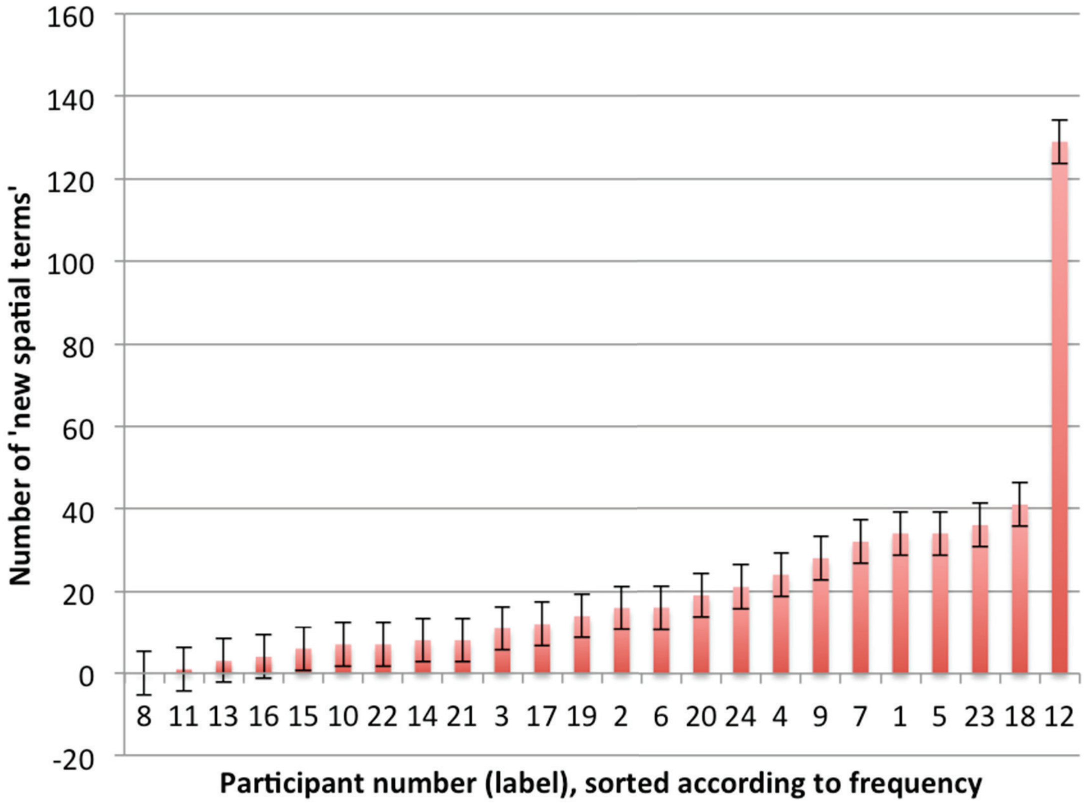


Figure 2. Number of "new spatial terms" used by individual participants, sorted according to frequency.

nouns and 10 "other" spatial terms occurred anywhere in the original instructions. This participant also produced the second highest amount of utterances categorised as "spatial description" (namely seven), including further interesting references to concepts such as a *kite like shape*, *valley vs. mountain folds*, *small flaps on the side,* and rationalisations such as *that little crease, that little corner, that obtuse angle right there, was meant to intersect with the centre line.*

Obviously, this particular participant was both highly verbal and outstandingly creative in range of spatial vocabulary. Other participants did show a similar kind of flexibility, albeit with lower frequencies of 'new' spatial terms (cf. Figure 2). Although the types of spatial descriptions produced by the participants varied considerably, these examples provide a representative impression of verbalisations when considering spatial instructions. The details and frequencies may differ, but the procedure appears to be comparable across those participants who produced "spatial descriptions" and "new spatial terms."

Verbalisations of problems were typically less explicit; participants said things like "that doesn't seem right" or "I was a little confused of this," without specifying further. More explicit statements in this category include "I wonder if that was still supposed to be folded somehow," "does the angle sort of matter?," "that's a little lopsided," and "I think is just opening it up, right?" Thus, participants wondered aloud (without interacting with the experimenter) about the precise action to be carried out or the degree of precision to be pursued, were unsatisfied with the product, or tried to interpret the formulation used in the instructions. Often enough, this included some degree of spatial term use as well (i.e., verbalisation of spatial thinking).

VERBALISATION PATTERNS

After having identified the content and significance of the verbalisation categories as just outlined, the next step was to address patterns of recurring thoughts or processes as reflected in the think-aloud data. For this purpose, we analyzed the frequency and distribution of the categories (ignoring the *Other* category, which was rarely used and contained unintelligible parts that did not lend themselves to counting) in relation to folding steps and participants, and determined the order of category mention within each step.

The category *Reading* was used most often (232 times, averaging 0.74 per participant and step), and *Background knowledge* least often (28 times, averaging 0.09). The other categories fell in between (*Reformulating*: 154 (0.49); *Additional ideas:* 165 (0.53); *Evaluation:* 49 (0.16); *Problems:* 105 (0.34); *Task communication:* 191 (0.61)). Participants used most categories frequently, though not necessarily for each step. All but two participants *read* parts of the instructions verbatim, all participants *reformulated* something at least once, and all but two formulated *additional ideas* about at least one step. Seventeen participants *evaluated* their own work, 20 *expressed problems*, and all except one *communicated about the task*. Only references to *background knowledge* were less frequent, occurring with only 8 participants (although 22 of 24 reported some previous Origami experience; see further details below). Figure 3 illustrates the frequency with which each participant used a category.

The distribution of the categories across the folding steps (i.e., associated to the steps shown in Appendix C) was informative. *Readings* and *reformulations* were fairly evenly distributed across the folding steps (ranging between 8 and 18 mentions by different participants in each of the 13 steps), indicating that the content of a folding step did not affect (in any obvious way) whether an instruction was read verbatim, reformulated, or just comprehended without verbalising. The other categories were not distributed equally. For instance, the instruction for folding step 4 was "Turn around 180 degrees," which never induced any *further ideas*, nor *evaluations* or references to *background knowledge*, and only one *expression of problems*. In contrast, folding step 6 was "Fold the bottom edges onto the midline once again," which led 15 (of the 24) participants to formulate *additional ideas* such as "so the same thing" (across-step repetition), "the same with the left" (within-step repetition), or "so in half again" (spatial description). The other categories were represented more frequently in other steps. Figure 4 visualises the distribution of category usage across folding steps.

To shed further light on differences between individual folding steps in terms of behaviourally reflected cognitive complexity, we calculated the mean number of words produced as well as the time needed for each folding step. Divided by 10 to match scale, the mean number of words is imposed within Figure 4 (dashed line) to reveal a clear visual effect: the number of words used along with a specific folding step generally matches the pattern of number of participants producing verbalisation types for the same step. The number of words peaks at folding step 3 (mean = 72.04 words produced) followed by folding step 10 (mean = 50.71); these are the steps for which most participants explicitly mentioned problems. The lowest number of words were produced along with folding steps 4 (mean = 11.46) and 7 (mean = 16.33). Both of these triggered few *problems* or *additional ideas*, and the like, as shown in Figure 4. The folding times needed for these steps matched this pattern, with a relatively high average fold time of 47 seconds for step 3 followed by 30 seconds for step 10. In contrast, the simpler folding step 4 required

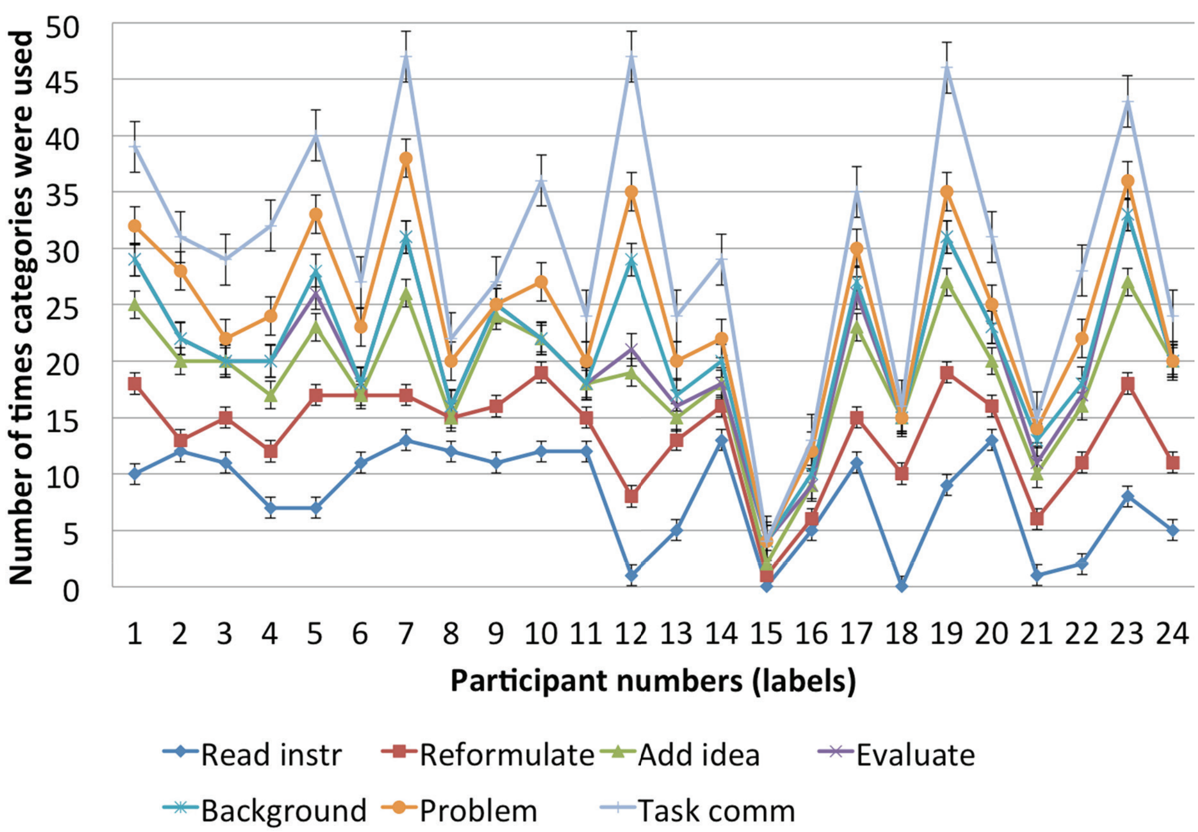


Figure 3. Stacked frequency of category usage across participants.

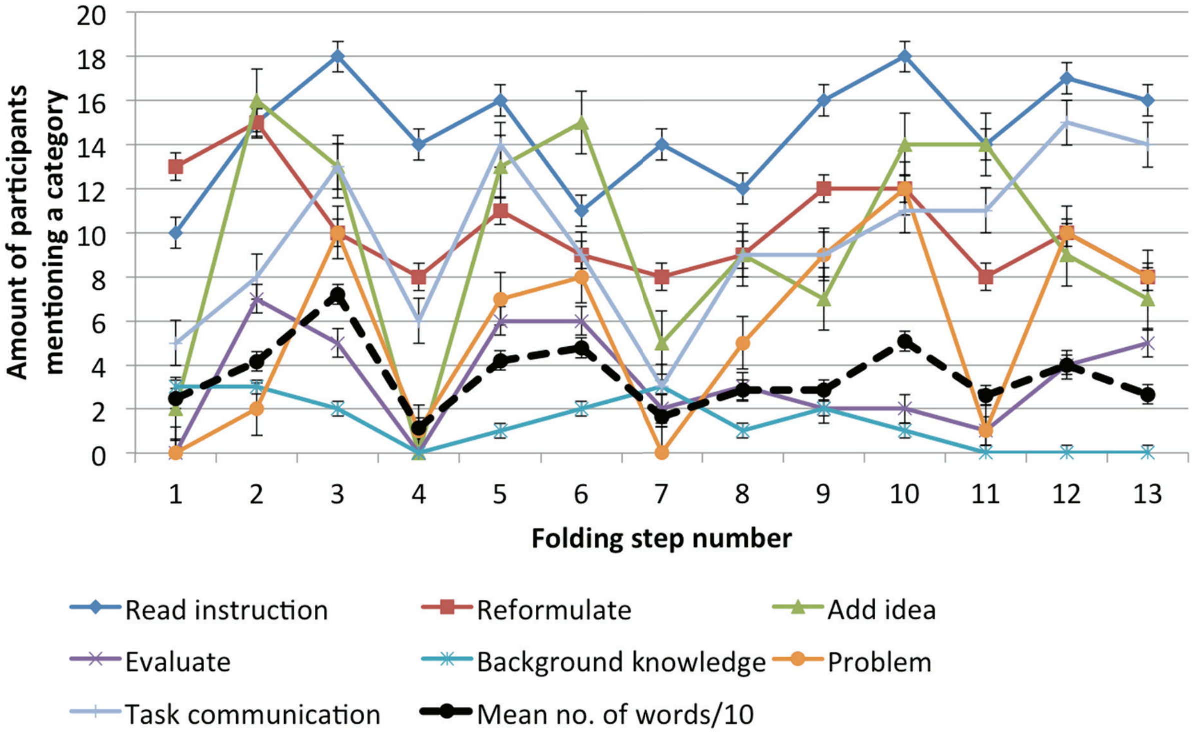


Figure 4. Amount of participants using a category in each folding step (1–13), and mean number of words used for each step (divided by 10 to match scale).

only 10 seconds on average, and step 7 took 13 seconds. This converging evidence points to systematically different levels of cognitive complexity as reflected by various behavioural measures, leading to differences in the verbalised problem solving processes.

In addition to differences between steps, there were also differences between individuals. For instance, three participants read verbatim parts of every step, while two never read. Based on the ratio of reading to reformulating, we identified 10 *readers*, 5 *reformulators*, and 9 participants who were *neutral* in this respect (ratio ranging between 0.7 and 1.83 or, in one case, producing only one reformulation and no readings). This indicates individual approaches to dealing with the original formulation and reverbalisation. Furthermore, one participant produced additional ideas for as many as 11 of the 13 steps, while most others did this far less frequently (overall mean = 5.17). Also, 8 participants never explicitly referred to the picture, while the others did this at least once and up to four times throughout the folding task (overall mean = 1.54).

Given this diversity in verbalising concepts and ideas relevant to the folding process, we looked for systematic patterns based on the order in which these ideas were mentioned, if they occurred at all within a folding step. While the above analysis merely established whether or not a category appeared in a folding step, a closer look revealed that they did not appear in random order, in spite of a high diversity of combinatorial possibilities. Consider the example in Table 2. After a discourse marker *alright* marking the start of the task, the participant reads step 1, with a slight grammatical reformulation at its end (*me* rather than *you*). This is followed by an action comment (*task communication*): *I'm doing that.* The phrase *next thing I need to* is again a comment on carrying out the task, introducing the reading of the next instruction. In the next line, "so it's lined up" expresses the additional concept of "alignment," not explicitly given in the instruction. This is followed by a slight reformulation of the next part of the instruction, "straighten out the fold." From here, the participant proceeds to the next step, which starts by reading the instruction and communicating about the task. The phrase *so I guess that means* indicates a certain concern about the correct interpretation, followed by a reformulation of the task (*to the center* rather than *midline*), without expressing a different idea. The next three utterances reflect the participant's conceptual development moving away from the original instruction; none of these are directly expressed in the instruction. While the instruction only uses a plural form ("edges"), the action itself needs to be carried out twice, which is expressed by *I gotta do it with the other side* (coded as *within-step repetition*).

Table 2.
Think aloud example, moving from reading to reconceptualising

Step	Category	Utterance
1	*read, reformulate*	alright. so put the paper in front of you with one corner pointing towards me
1	*task communication*	I'm doing that
2	*task communication, read*	next thing I need to fold the left corner over to the right one
2	*task communication, additional idea, reformulate*	so I'm folding that so it's lined up and then straightening out the fold
3	*task communication, read*	so it's telling me to open the paper again
3	*task communication, read*	so I do that and fold the bottom edges towards the midline
3	*express problem, reformulate*	so I guess that means fold this part to the center here
3	*additional idea*	and I'm trying to do that so it's as even as possible
3	*additional idea*	now I gotta do it with the other side
3	*additional idea*	and crease that part there

Table 3.
Think aloud example involving expression of problems

Step	Category	Utterance
3	*read*	uh, open the paper again
3	*read, reformulate*	fold the bottom edges towards the middle
3	*express problem*	okay, um, how should I do that?
3	*read*	uh, open the paper and fold the bottom edges toward the midline
3	*express problem*	but how did it turn out like that?
3	*express problem*	that's kinda confusing; also kind of annoying
3	*other*	so, do do do
3	*reformulate*	fold open

In contrast, the Table 3 example reflects an inability to move away from the original instruction. The task instruction is repeatedly read, interspersed by expressions of confusion.

A close scrutiny of the linguistic data led to the identification of a basic recurring pattern as follows:

reading—reformulating—reconceptualising—evaluating

Participants started out by *reading* a folding step, (possibly) followed by *reformulations*. Conceptually moving away from the original wording, they would then (in more fortunate cases than the one shown in Table 3) verbalise *additional ideas* about this folding step, and then possibly *evaluate* their product. Although past *experience* could be verbalised at any point in the process, it typically appeared only after evaluation (if any). Although folding actions are not reflected directly in the verbal data and were not annotated in this study, participants sometimes referred to action (e.g., "which I did here"), and this was categorised as *task communication*. This emerged as a free-floating category, appearing anywhere in the process (including leading over to next step). This reflects how participants were continuously acting on their object (as expected), following the instructions as soon as they were able to interpret them, based on comprehension and reconceptualisation processes. *Problems* led to disruptions of this process, with participants either unable to move away from the instruction at all as in the above example, or starting again by reading, or anywhere else within the overall process.

To verify this intuition, we identified for all utterances made by participants in relation to individual folding steps whether or not they were consistent with the pattern *reading—reformulating—reconceptualising—evaluating*, treating *expression of problem* as a reset to start, and ignoring *task communication* and *other*. Steps could be skipped, since think-aloud protocols can never be expected to represent all thought processes exhaustively). A paired *t*-test (comparing consistent vs. inconsistent patterns within each folding step) showed that the verbalisations within steps were consistent with this overall process scheme significantly more often than they were inconsistent ($p < .001$). However, even within-step verbalisations that did not precisely fit this pattern seemed close enough to support the overall scheme, as in the Table 2 example (*so I'm folding that so it's lined up and then straightening out the fold*): Here, the process of reading was interrupted as a partial step triggered some thoughts (here: *alignment of the folds*, as a type of *additional idea*). Nevertheless, the overall scheme ranging from mere reading via slightly reformulating to reconceptualising and evaluating (and possibly reminiscing previous experiences) was generally followed, in a flexible way.

EXPERIENCE

Since previous work on problem solving consistently showed effects of background knowledge and previous experience, affecting not only overall performance but also problem solving strategies and pathways, we finally addressed how reported Origami experience related to verbalisation in the given task. Of our 24 participants, when asked about their background, 19 reported having some previous experience with Origami, 3 a lot, and 2 none. In spite of the somewhat uneven distribution in this regard, we were able to detect some interesting associations using Pearson's correlation. Experience was positively correlated with overall verbosity, that is, number of words produced ($r = 0.48$, $n = 24$, $p < .05$); this corresponded to a higher number of units ($r = 0.37$, $n = 24$, $p < .05$) as well as a higher number of words per unit ($r = 0.48$, $n = 24$, $p < .05$). Looking at content, it turned out that the verbalisations were more creative with more experience; experience was marginally negatively correlated with the number of words read ($r = -0.35$, $n = 24$, $p = .091$), but positively with the number of words expressing additional ideas ($r = 0.38$, $n = 24$, $p = .067$). More experienced Origami users carefully compared their work with the instruction often ($r = 0.56$, $n = 24$, $p < .01$), and they used more new spatial terms to verbalise their thoughts ($r = 0.48$, $n = 24$, $p < .05$). Also, they reliably used more words to communicate about the task ($r = 0.56$, $n = 24$, $p < .01$), and they (expectedly) talked more about past experience ($r = 0.58$, $n = 24$, $p < .01$). Along these lines, previous experience affected how people distributed their verbalisations, without fundamentally changing the overall pattern (as there were no outliers with respect to any of our analysis categories).

DISCUSSION

Most problem solving research focuses on unaided tasks, reflecting an implicit assumption that instructions guide cognitive processes sufficiently to leave few or no problems to be solved. Generally, the extant literature suggests that step-by-step instructions *lead* rather than *trigger* trains of thought; complete instructions should leave little room for creativity. Accordingly, think-aloud protocols when following such instructions should not contain much content beyond a reflection of the guidance the instructions provide. However, because findings on reading and visual comprehension point to a more complex process when interpreting action instructions, our study set out to challenge this assumption. We used a task that followed established traditions in the Japanese art of Origami folding. It used step-by-step instructions that were complete in the sense of guiding the reader through the whole process from a blank piece of paper to the completed product, without omitting any actions. Indeed, none of our participants mentioned a need for further instructions; any problems that were expressed had to do with the actions involved *within* an instruction step. The guidance was complete at the overall task's highest level. Nevertheless, it left room for interpretation, highlighting a different layer of problem solving processes. Our results, drawn from verbalisations uttered throughout the task, point to distinct cognitive processes involved in understanding and completing the instructions.

Notably, the level of problem solving we see here does not correspond, as might have been assumed, to an automated subconscious level of task execution. Instead, our results suggest interpretation processes that are consciously accessible and verbalisable to a high extent, even where the main solution steps are available in both verbal and pictorial format.

Our results speak to a range of findings across domains such as general problem solving, instruction interpretation, spatial reasoning, discourse comprehension, educational practices, and verbalisation of thought. We will address each of these in turn.

Research on problem solving, in general, typically aims to identify the main solution steps (sub-goals) and cognitive strategies employed commonly by humans solving complex problems (following Newell & Simon, 1972). Once these have been determined, for example based on verbal protocols (Ericsson & Simon, 1984), they can be represented in terms of computational models and cognitive architectures (Anderson, 2004; Gugerty & Rodes, 2007; Pizlo et al., 2006). Beyond high-level cognitive operators, such models also include more fine-grained representations of how the action steps are accomplished. However, typically these are not expressed in terms of problem solving processes as such.

Our research suggests a different picture. Apparently, specifying the main solution steps in a complete set of step-by-step instructions does not eliminate the need for problem solving. Instead, the main solution steps provide a coarse level of problem solving, but leave room for more fine-grained challenges. Our verbal data highlight the cognitive complexity and creativity involved in this process, going well beyond the immediate and automatic interpretation of clearly laid out instructions. In our data, verbalisations coded as *Additional ideas* signaled how the participants went beyond the original instruction, that is, the surface of the formulation. This was a frequent category in our data, used in about half of the folding steps by each participant; with two exceptions, all participants did this at least once. Therefore, the reconceptualisation of the original instruction by adding related ideas played a major role within the procedural pattern detected in our data.

However, such reconceptualisation was not engaged within every instruction step to the same extent. Very simple instructions (e.g., *turn paper around)* could be directly transformed to action and did not appear to trigger further thoughts. With increasing task complexity, participants took consistently more time to accomplish and words to verbalise an instruction step. These reconceptualisations highlight the active consideration of how to appropriately follow the instructions. Thus, rather than simply following the instructions verbatim and activating automated processes, participants actively engaged in thought processes, going well beyond the step-by-step guidance given to them. In line with previous literature on cognitive complexity (Alter et al., 2007; Garrod, 1985), this effect was mediated by the level of challenge posed by any individual action step. More complex instructions led to more verbalised thoughts and inferences, as well as increased expression of problems.

The literature pointing to the existence of complex interpretation processes and inferences involved in reading and pictorial comprehension (Hegarty & Just, 1993; Kosslyn, 1980; Zwaan & Radvansky, 1998) corroborates our results. These processes take the reader well beyond a message's surface representation (Bransford & Franks, 1972; Franck, 2004) and trigger cognitive effort at levels similar to other problem solving processes (Alter et al., 2007; Garrod, 1985). Some of these processes include selective attention (Franconeri et al., 2012) guided by relevance (Florax & Ploetzner, 2010) and background knowledge (Glenberg & Robertson, 1999). In fact, the necessity of activating inference processes may be beneficial for deep understanding (McNamara et al., 1996). However, coming from research unrelated to problem solving, this work does not reveal any particular sequence of interpretation processes. Our analysis of thoughts verbalised during Origami paper folding sheds more light on these issues in the form of recurring patterns in the language data. Participants gradually moved away from the original wording toward a reconceptualisation of an instructional step. Starting out by reading parts of the instructions verbatim, they quickly turned to minor reformulations, then added additional ideas, before evaluating their product. To our knowledge, our approach provides the first operationalisation for the systematic analysis of verbalisations related to an instruction interpretation process (including reading and picture comprehension and transfer toward real world action).

Task reconceptualisation, as reflected in our annotation category of *Additional ideas*, may be viewed as a verbal representation of a cognitive process essential to Origami paper folding—namely, transferring the abstract textual content (supported by a 2-D picture) to concrete action. To do this, people need to understand the instruction and (creatively) interpret (or, indeed, reconceptualise) it in relation to their own product—going well beyond a direct or (nearly) automatic transfer from readily laid out operations that leave no room for problem solving. In some cases, they formulate specific additional ideas that are particularly clear in their own minds. In other cases, the interpretation and reconceptualisation processes may not be verbalised even if they do occur on some, perhaps less consciously accessible, level. Generally, people do not explicitly formulate the transfer process when thinking aloud (e.g., by saying "I am now trying to transfer this instruction to the piece of paper in my hand")—this would be easily accessible through content analysis (Krippendorff, 2004). Instead, the present study gained insights into participants' thoughts through a close analysis of the language data. This analysis revealed insights beyond the content of the explicit verbalisations by identifying utterance types relative to the instruction, and by analysing spatial term use. This approach is in line with Cognitive Discourse Analysis (CODA, Tenbrink, 2015), with its main goal of interpreting language use in relation to thought. The present analysis highlighted thought processes during instruction interpretation, and led to further insights about the role of verbalisation. As suggested by Taylor and Tenbrink (2013), access to relevant vocabulary for an idea can be helpful when implementing that idea on subsequent tasks. Another striking aspect of the reconceptualisation, as observed here, is the fact that many participants actually volunteered revised spatial descriptions, associating various concepts and spatial relationships. This suggests that participants actively sought to thoroughly understand the spatial situation, and expressed their own representation beyond the one provided.

The idea that people transform and reconceptualise a description in relation to the real world situation at hand resonates with findings in other areas of spatial discourse. According to Tenbrink et al. (2008), recipients of spatial instructions frequently provide insightful ideas that complement the verbal instruction given to them, filling in conceptual gaps using available perceptual information. More generally, interpreting spatial language inevitably depends on intricate inference processes that may involve drawing on background knowledge and judgments about the speaker providing the description (Gagnon et al., 2012; Gondorf, Bergmann, & Tenbrink, 2012). In this light, adding one's own ideas while interpreting an instruction seems only natural, since a direct mapping of linguistic descriptions to real-world objects is rarely possible. In route instructions, for instance, the main route is laid out in the descriptions, but there is still potential to go wrong in the real world for many reasons, including miscommunication, memory failure, reference resolution problems, underspecification, false information, and perspective and orientation problems. In short, following instructions in a spatial setting introduces a range of problems to solve, necessitating creative and active thought processes such as those reflected in our think-aloud protocols. To our knowledge, our study is the first to outline these phenomena in detail based on language data analysis.

Another frequent and everyday observation relevant to our findings is that, upon receiving a complex explanation in a face-to-face situation, it may be perceived as insufficient to simply acknowledge the information by nodding or responding "OK." Such feedback may be due to (or attributed to) politeness rather than true understanding. Arguably, the more complex an instruction or explanation is, the more reconceptualisation will be needed to demonstrate deeper understanding. This is particularly pertinent in school education. Teachers actively elicit summaries and reformulations on a regular basis; as written text types, they are integral parts of teaching approaches. Being able to summarise and reformulate is thus a skill to be learned because it can demonstrate comprehension that goes beyond the input itself (e.g., Chi, De Leeuw, Chiu, & Lavancher, 1994).

While reconceptualising and reformulating upon request by a teacher may be a cognitive challenge, it may also support the learning process. Verbalisation and access to associated terminology can support cognition, as demonstrated by research in two directions. First, various studies have indicated an enhancement of problem solving processes via verbalisations while doing the task. Fairly uncontroversially, providing good and elaborate explanations while studying examples correlates with success in problem solving (Chi et al., 1989); being asked for explanations and background information supports depth of thought and therefore enhances the problem solving process (Bielaczyc, Pirolli, & Brown, 1995; Neuman & Schwarz, 1998). However, whether or not simply *thinking aloud*—rather than providing explanations—serves to support cognitive processes appears to be dependent on the problem solving task and the way the instruction is formulated (Ericsson & Simon, 1984; Fox, Ericsson, & Best, 2011; Schooler et al., 1993).

The present study was designed to address the *nature* of verbalisable cognitive processes rather than their *effects* on performance (with performance measures only affecting minor parts of our analysis); a control group without verbalisation would not have led to any insights in this regard. However, based on the insights gained here, an informative next study could explore whether thinking aloud helps participants accomplish the Origami tasks successfully. In our study, we did not find correlations between success in the Origami paper folding task and reconceptualisations in the verbal data. However, as discussed in Taylor and Tenbrink (2013), use of new spatial terms was correlated with another measure, namely performance in the crease-pattern matching task given to participants after the folding task. This indicates that creative verbalisation can relate to performance in spatial tasks in somewhat intricate ways. The ability to verbalise spatial relations may enhance spatial thinking in a general sense, even if it does not directly affect the currently verbalised task.

The second research direction relevant to the cognitive effects of verbalisation addresses the relationship between language and thought, as critically inspired by Whorf (1941). In particular, evidence is accumulating that inner speech and labeling systematically support cognition at various levels, ranging from perception to categorisation and memory (Lupyan, 2012). Linguistic formulation of perceived categories appears to support ongoing conceptualisations by capturing fleeting impressions in a temporary way, supported further by previous linguistic experience and knowledge. It appears that fairly similar processes may be at work in our Origami task, in spite of the fact that labels exist, through the instructions. Our participants made heavy use of these existing formulations by reading aloud and modifying them only slightly at first, but then moved on to new conceptualisations and associating linguistic labels with them. Clearly, since they were not asked to formulate anything in particular (just *think aloud*), they chose descriptions relevant for them (i.e., they found labels and highlighted spatial relationships as they became obvious in their minds). Thus, while the research reported here was not designed to test whether reformulations and reconceptualisations actually support the problem solving process, our empirical findings do show that this cognitive process is an integral part of a cognitive task that is considerably more complex than the labeling of a newly encountered object.

CONCLUSION

Our study provides insights about the cognitive processes involved in following Origami paper folding instructions, challenging the assumption that following instructions leads to straightforward action execution. Instead, problem solving can be viewed as a multilayered process—not only in terms of high-level (conscious) and low-level (automated) processes, but also in terms of main problem solving steps (provided in complete instructions) and intermediate problems needing to be solved to accomplish these main steps. This level involves both high-level and low-level cognitive processes and is therefore in part explicitly verbalisable, and in part reflected in the linguistic features and patterns of the verbalised data.

Our results suggest a recurring pattern of gradually moving away from the original instruction by reading, reformulating, adding ideas and associated concepts, and evaluating the folding effort (with a possible addition of background experience). This pattern highlights the necessary conceptual path involved in interpreting an abstract instruction in such a way as to act appropriately in the real world. Specifically, it supports the theory that reconceptualisation—be it through explicit verbalisations, or only silently in the mind—is an important and supportive part of this comprehension process. Further research is required to explore the extent to which explicit verbalisation introducing new formulations can support problem solving processes.

REFERENCES

Afflerbach, P., & Johnston, P. (1984). On the use of verbal reports in reading research. *Journal of Reading Behavior, 16,* 307–322.

Allen, G. L. (2000). Principles and practices for communicating route knowledge. *Applied Cognitive Psychology, 14,* 333–359.

Anderson, J. R. (2004). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060.

Anderson, J. R., Douglass, S., & Qin, Y. (2004). How should a theory of learning and cognition inform instruction? In A. Healy (Ed.), *Experimental cognitive psychology and its applications*. Washington, D. C.: American Psychological Association.

Anderson, J. R., Kline, P., & Lewis, C. (1977). A production system model for language processing. In P. Carpenter & M. Just (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Lawrence Erlbaum.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136,* 569–576.

Ball, J. (2004). A cognitively plausible model of language comprehension. In *Proceedings of the 13th Conference on Behavior Representation in Modeling and Simulation* (pp. 305–316). Red Hook, NY: Curran Associates, Inc.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–609.

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review, 26*(1), 27–50. http://dx.doi.org/10.1007/s10648-014-9251-9

Bergen, B., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science, 31,* 733–764.

Bielaczyc, K., Pirolli, P., & Brown, A. (1995). Training in self-explanation and self-regulation strategies: investigating the effect of knowledge acquisition activities on problem-solving. *Cognition and Instruction, 13,* 221–253.

Bransford, J. D., & Franks, J. J. (1972). The abstraction of linguistic ideas: A review. *Cognition, 1*(2–3), 211–249.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Carruthers, S., & Stege, U. (2013). On evaluating human problem solving of computationally hard problems. *The Journal of Problem Solving, 5*(2), 42–70.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication.* Oxford, UK: Blackwell.

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

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439–477.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Hillsdale, NJ: Erlbaum.

Cohen, C. A., & Hegarty, M. (2012). Inferring cross sections of 3D objects: A new spatial thinking test. *Learning and Individual Differences, 22*(6), 868–874.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90, 1073–1091.

Daniel, M.-P., & Tversky, B. (2012). How to put things together. *Cognitive Processing, 13*(4), 303–319.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis—Verbal reports as data*. Cambridge, MA: Bradford.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *The Journal of Problem Solving, 4*(1), Article 3. http://dx.doi.org/10.7771/1932-6246.1118

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction, 20,* 216–224.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137,* 316–344.

Franconeri, S. L., Scimeca, J. M., Roth, J. C., Helseth, S. A., & Kahn, L. (2012). Flexible visual processing of spatial relationships. *Cognition, 122*(2), 210–227.

Furuyama, N. (2000). Gestural interaction between the instructor and the learner in origami instruction. In D. McNeill (Ed.), *Language and gesture* (pp. 99–117). Cambridge, UK: Cambridge University Press.

Gagnon, S. A., Brunyé, T. T., Tenbrink, T., Gopal, N., Gardony, A. L., Holcomb, P. J., & Taylor, H. A. (2012). To err is human: Landmark vs. turn reliance under conditions of route ambiguity. *53rd Annual Meeting of the Psychonomic Society,* November 15–18, Minneapolis, MN.

Garrod, S. C. (1985). Incremental pragmatic interpretation versus occasional inferencing during fluent reading. In G. Rickheit & H. Strohner (Eds.), *Inferences in text processing* (pp. 161–181). Amsterdam: Elsevier.

Gero, J. S., & McNeill, T. (1998). An approach to the analysis of design protocols. *Design Studies, 19*(1), 21–61.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes, 28*(1), 1–26.

Gondorf, C., Bergmann, E., & Tenbrink, T. (2012). Spatial inferences and language use. *4th UK Cognitive Linguistics Conference*, July 10–12, Franklin Wilkins Building, Kings College London.

Gugerty, L., & Rodes, W. (2007). A cognitive model of strategies for cardinal direction judgments. *Spatial Cognition and Computation, 7*(2), 179–212.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32,* 717–742.

Henderson, J. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504.

Higginson, W., & Colgan, L. (2001). Algebraic thinking through origami. *Mathematics Teaching in Middle School, 6*(6), 343–350.

Hölscher, C., Büchner, S.J., Meilinger, T., & Strube, G. (2009). Adaptivity of wayfinding strategies in a multi-building ensemble: The effects of spatial structure, task requirements, and metric information. *Journal of Environmental Psychology, 29*(2), 208–219.

Hölscher, C., Tenbrink, T., & Wiener, J. (2011). Would you follow your own route description? *Cognition, 121,* 228–247.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research, 13,* 203–212.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Krause, P., Reyle, U., & Schiehlen, M. (2001). Spatial inferences in a localization dialogue. In M. Bras & L. Vieu (Eds.), *Semantic and pragmatic issues in discourse and dialogue* (pp. 183–216). Amsterdam: Elsevier.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). London, UK: Sage.

Kuipers, B. J., Moskowitz, A. J., & Kassirer, J. P. (1988). Critical decisions under uncertainty: Representation and structure. *Cognitive Science, 12,* 177–210.

Larkin, J. (1989). Display based problem solving. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 319–341). Hillsdale, NJ: Erlbaum.

Lee, P. U., & Tversky, B. (2005). Interplay between visual and spatial: The effects of landmark descriptions on comprehension of route/survey descriptions. *Spatial Cognition and Computation, 5,* 163–185.

Lovelace, K. L., Hegarty, M., & Montello, D. R. (1999). Elements of good route directions in familiar and unfamiliar environments. In C. Freksa & D. M. Mark (Eds.), *Spatial information theory* (pp. 65–82). Berlin: Springer.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Cognition, 3*(54). http://dx.doi.org/10.3389/fpsyg.2012.00054

Marcus, N., Cooper, M. G., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology, 88*(1), 49–63.

McNamara, D. S., Kintsch, E., Butler Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14(*1), 1–43.

Meilinger, T., & Knauff, M. (2008). Ask for your way or use a map: A field experiment on spatial orientation and wayfinding in an urban environment. *Spatial Science, 53*(2), 13–24.

Muller, P., & Prévot, L. (2009). Grounding information in route explanation dialogues. In K. Coventry, T. Tenbrink, & J. Bateman (Eds.), *Spatial Language and Dialogue* (pp. 166–176). Oxford, UK: Oxford University Press.

Myers, C., Gluck, K. A., Gunzelmann, G., & Krusmark, M. (2010). Validating computational cognitive process models across multiple timescales. *Journal of Artificial General Intelligence, 2*(2), 108–127.

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

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

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition, 28,* 39–58.

Pizlo, Z., Stefanov, E., Saalwaechter, J., Li, Z., Haxhimusa, Y., & Kropatsch, W. G. (2006). Traveling salesman problem: A foveating pyramid model. *Journal of Problem Solving, 1*(1), 83–101.

Ritter, F. E., & Larkin, J. H. (1994). Developing process models as summaries of HCI action sequences. *Human-Computer Interaction, 9,* 345–383.

Robichaux, R., & Rodrigue, P. (2003). Using origami to promote geometric communication. *Mathematics Teaching in Middle School, 9*(4), 222–229.

Sabbah, D. (1985). Computing with connections in visual recognition of Origami objects. *Cognitive Science, 9,* 25–50.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics, 2*(9), 437–442.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101–120.

Schnotz, W., Baadte, C., Johnson, A., & Mengelkamp, C. (2012). Knowledge acquisition from verbal and pictorial information. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning* (pp. 339–365). Cambridge, UK: Cambridge University Press.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122*(2), 166–183.

Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology, 3*(2), 228–243.

Shimanuki, H., Kato, J., & Watanabe, T. (2003). Recognition of folding process from origami drill books. In *Proceedings of 7th International Conference on Document Analysis and Recognition* (pp. 500–554), August 3–6, Edinburgh, UK. http://dx.doi.org/10.1109/ICDAR.2003.1227725

Taylor, H. A., & Hutton, A. (2013). Think3d! Training spatial thinking fundamental to STEM education. *Cognition & Instruction, 31*(4), 434–455.

Taylor, H. A., & Tenbrink, T. (2013). The spatial thinking of Origami: Evidence from think-aloud protocols. *Cognitive Processing, 14,* 189–191.

Tenbrink, Thora. 2015. Cognitive Discourse Analysis: Accessing cognitive representations and processes through language data. *Language and Cognition, 7*(1), 98–137.

Tenbrink, T., Andonova, E., & Coventry, K. (2008). Negotiating spatial relationships in dialogue: The role of the addressee. In J. Ginzburg, P. Healey, & Y. Sato (Eds.), *Proceedings of LONDIAL—The 12th SEMDIAL workshop*, King's College, London, UK.

Tenbrink, T., & Wiener, J. (2009). The verbalization of multiple strategies in a variant of the traveling salesperson problem. *Cognitive Processing, 10*(2), 143–161.

Tversky, B. (2011). Visualizations of thought. *Topics in Cognitive Science, 3,* 499–535.

Van Deemter, K., & Peters, S. (Eds.). (1996). *Semantic ambiguity and underspecification*. Stanford, CA: CSLI Publications.

Van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*– 237–244.

Weiß, P., Pfeiffer, T., Eikmeyer, H.-J., & Rickheit, G. (2006). Processing instructions. In G. Rickheit & I. Wachsmuth (Eds.), *Situated communication* (pp. 31–76). Berlin: De Gruyter Mouton.

Whorf, B. L. (1941). The relation of habitual thought and behavior to language. In L. Spier (Ed.), *Language, culture, and personality: Essays in memory of Edward Sapir* (pp. 75–93). Menasha, WI.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation Models in Language Comprehension and Memory. *Psychological Bulletin, 123*(2), 162–185.

APPENDIX A

Instructions given to the participant to explain and train how to *think aloud*, following Ericsson and Simon (1984):

> "I will, in a minute, give you a task to perform. While you do that, I will ask you to THINK ALOUD during the whole procedure of the task. We are interested in what you think about as you perform the task. Therefore I want you to say EVERYTHING you are thinking from start to finish of the task. Don't try to plan out what you say and don't talk to ME. Just act as if you were speaking to yourself. It is most important that you keep talking, even though you won't get any response or feedback. Do you understand what I want you to do? If I do not hear you talking for a bit, I will remind you that you are to say aloud what you are thinking.
>
> Good, now we will begin with some practice problems. First, I want you to multiply two numbers in your head and speak out loud what you are thinking as you get an answer.
>
> What is the result of multiplying 24 x 36?
>
> Good. Any questions? -- Here's your next practice problem:
>
> How many windows are there in a house you used to live in—for example your parents' house?"

APPENDIX B

Instruction for Task 1—Origami paper folding

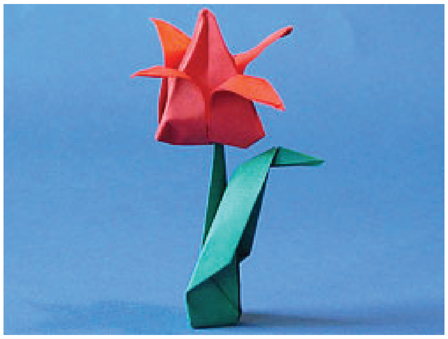
> "Okay, we are now ready to start with your first task. Here [show participant the instruction on the screen] is an instruction for an Origami paper folding task. The aim is to create an Origami paper tulip made of two pieces of paper, following these instructions. Start with the STEM, which is easier. Don't forget to THINK ALOUD while doing so.
>
> Take as long as you like. I won't interrupt you, and I won't judge what you have done. We are interested in your thoughts while you do the paper folding.
>
> Okay? You can start right away. When you're done with the stem, proceed directly with the bloom."

APPENDIX C

Illustrated Origami instructions for the flower stem

Instructions are identical to those used in the study except the numbering, which was added here for purposes of referencing in this paper.

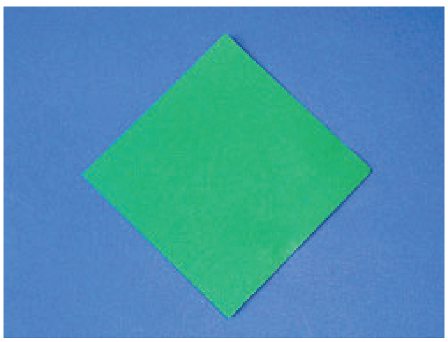
This is how you can make the stem



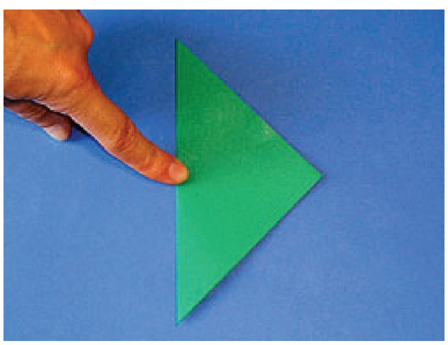
You'll need:

- Square green paper

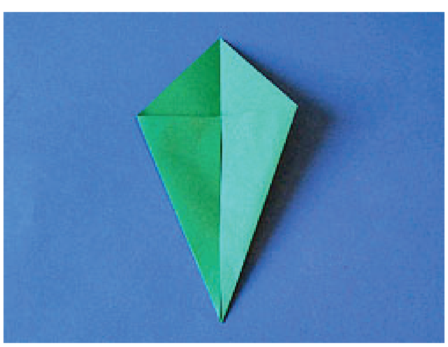
Instructions for the stem



1. Put the paper in front of you, with one corner pointing towards you.

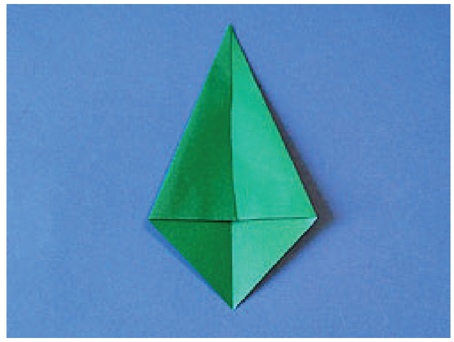


2. Fold the left corner over to the right one, and firmly straighten out the fold.

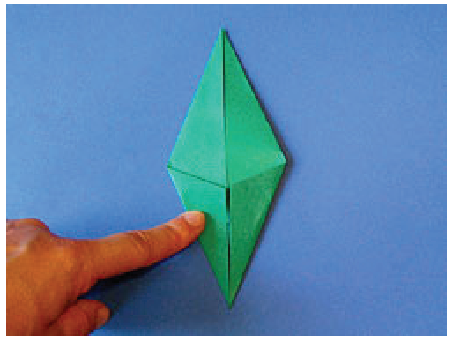


3. Open the paper again.

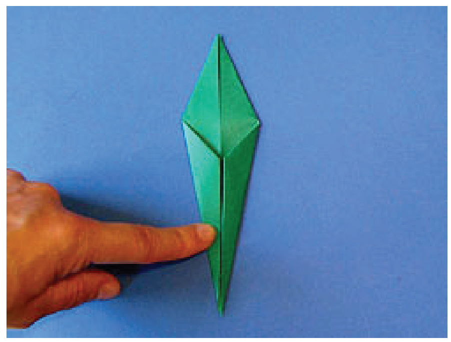
   Fold the bottom edges toward the midline.



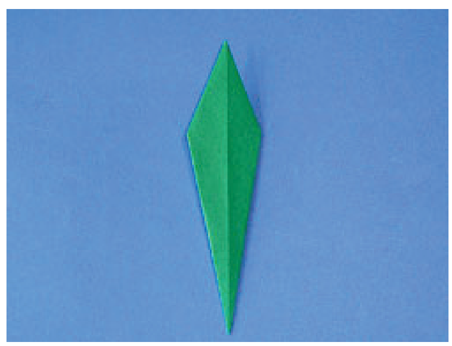
4. Turn around 180 degrees



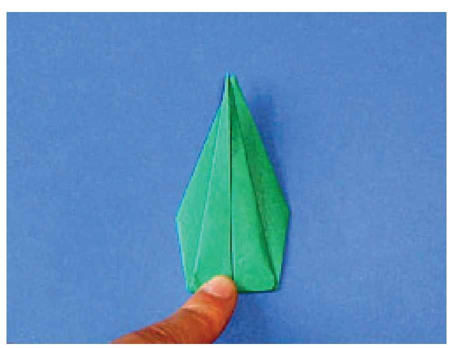
5. Fold the bottom edges onto the midline.



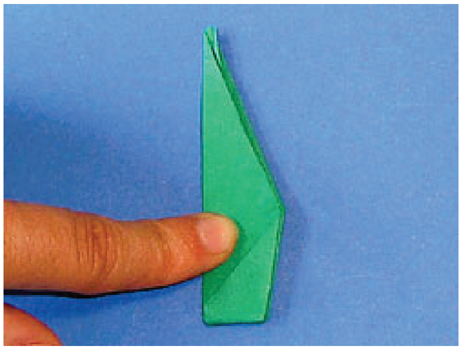
6. Fold the bottom edges onto midline once again.



7. Turn the paper over so that the back side is up.

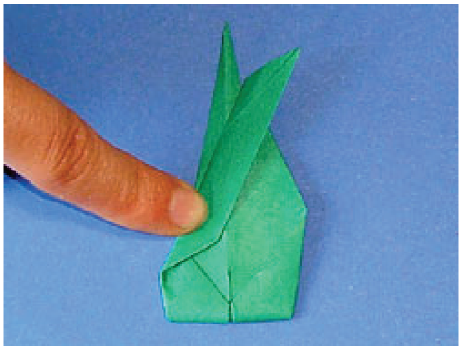


8. Fold the lower tip onto the upper one.



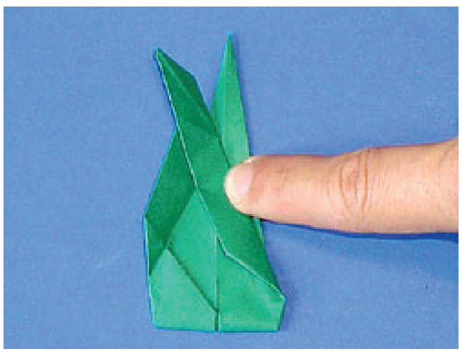
9. Fold the left half over the right one.

   Fold it back again.

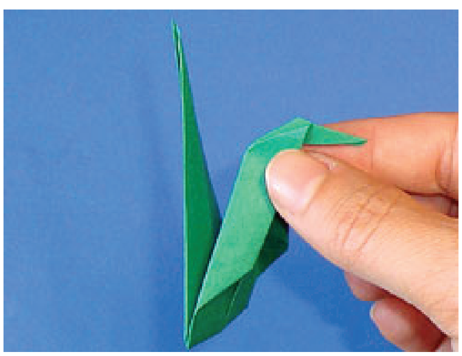


10. Turn the paper over.

    Fold the overlying layer of the upper tip diagonally to the right (kind of like in the picture)

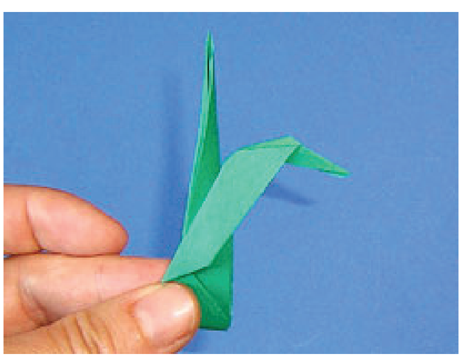


11. Fold it back, and then diagonally to the left.



12. Fold the tip back.

    Fold the upper tip toward the inside along the fold lines.



13. Push the left and the right half of the stem together.

APPENDIX D

Remaining new spatial terms (complementing Figure 1) sorted according to the absolute number of occurrence.

9	8	7	6	5	4	3	2	1	1	1	1
top	back	even	edges	along	edge	across	around	apart	face	loose	solid
	inside	right	left	diagonal	flat	angle	back	at-the-top	farther	lopsided	square
		sides	lines	flap	folds	bottom	center	balloon-ish	fit	lower	squares
		there	square	open-up	go	center	close	base	fits-in	match	staggered
				tip	in-half	creases	diagonal	baseline	fitting	middle-line	starting-point
					little	direction	far	baselines	flip-over	mirror	step
					open	end	flip	bent	fold-lines	narrow	straighten
					shape	flaps	flipped-over	big	folded	ninety-degrees	stuck
					symmetrical	fold	hole	bigger	folded-part	one-hundred-eighty-degrees	switch
					thick	from	inside	bisect	folding	opening	symmetry
					top	halfway	layer	bisect	forty-five-degrees	openings	thicker
					triangle	layers	lined-up	bisectors	from-the-top	opposed	three-dimension
					upper	middle	match	box	gap	opposed-to	three-dimensional
						onto	midline	bulgier	half	outside	tight
						opposite	off	came-out	halfway	overlapping	tighter
						straighten-out	open	center-line	higher	overlying	tips
						turn	opened-up	closest	hundred-eighty-degrees	parallelogram	to-the-half
						where	opening-up	collapsing	in-front	patterns	touching
							parallel	come	in-the-back	place	towards
							position	coming-out	inflate	pointing-out	tucked-in
							small	cone-shape	inner	portion	turn-around
							smooth	contour	interlace	preliminary	turn-over
							square-ish	creased	intersect	proportion	turned-around
							sticking-out	creasing	inverted	pyramid	two-dimensional
							through	cross	isosceles	radial	under
							touch	crumpled	lateral	reference-point	underlying
							toward	curling	line-up	reverse	underneath
							uneven	curve	lined	ripped	up-to
							upside-down	diamond	lining-up	seam	upperlying
							vertical	end-up	lip	size	upside
							overlap	expand	longer	smushed	visuospatial