

“Now move like that fish”: Can enactment help learners come to understand dynamic motion presented in photographs and videos?

Katharina Scheiter^{a,b*}, Birgit Brucker^a, Shaaron Ainsworth^c

^aLeibniz-Institut für Wissensmedien, Tübingen, Germany

^bUniversity of Tübingen, Germany

^cSchool of Education, University of Nottingham, England

*Correspondence concerning this article should be addressed to Katharina Scheiter, Leibniz-Institut für Wissensmedien, Schleichstraße 6, 72076 Tuebingen, Germany; email: k.scheiter@iwm-tuebingen.de.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Running head: Enacting Fish Locomotion

“Now move like that fish”: Can enactment help learners come to understand dynamic motion
presented in photographs and videos?

1
2
3
4 **“Now move like that fish”**: Can enactment help learners come to understand dynamic
5
6 **motion presented in photographs and videos?**
7
8
9

10
11 **Abstract**
12

13
14 Technological advancements offer new possibilities of interacting with learning
15 materials, including the use of gestures and body movements. The present study addressed the
16 question of how using one’s body to enact movements whilst learning about them would affect
17 outcomes. 85 participants were shown either sequences of photographs or videos of fishes
18 deploying different locomotion patterns for propulsion. Half of the participants in each
19 visualization condition were prompted to enact the movements whilst learning. During learning,
20 all participants were asked to rate their mental effort; moreover, their enactments were
21 videotaped and later coded with respect to their frequency of occurrence and their congruency
22 with the actual fish locomotion. After the learning phase, students were asked to classify fish
23 based on their locomotion behavior as well as to describe fish showing familiar and unfamiliar
24 locomotion behaviors to assess learning outcomes. Results showed that – independent of
25 visualization format – being asked to enact the fish movements had a positive effect on the
26 students’ ability to classify fish as long as the locomotion behavior was neither too easy nor too
27 difficult to be recognized. It did not affect the ability to describe fish movements. The frequency
28 of enactments and their congruency were unrelated to learning outcomes. Taken together, the
29 effects of enactments in this study appear to be limited to certain tasks. Furthermore, they are
30 likely to be due to enhancing engagement in the learning rather than to mechanisms specific to
31 enacting body movements.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Keywords: visualization; enactment; gestures; learning about movements

1. Introduction

The way people interact with learning technologies is changing. Whilst the keyboard and the mouse may still predominate, there are increasing applications which ask learners to write or draw with digital pens, feel forces in their fingers via haptic joysticks or manipulate physical objects to change their digital representation (tangibles). In particular, there are many more applications that allow people to control them with their hand and body movements, and such gesture-based learning environments are increasing rapidly (Sheu & Chen, 2014). This evolution in the technological landscape goes hand in hand with our growing knowledge of how learning with our bodies shapes our understanding (Alibali & Nathan, 2012; Glenberg, 2010; Goldin-Meadow & Wagner, 2005) and can be utilized to help people learn (Lindgren & Johnson-Glenberg, 2013). However, as yet there is more excitement than there are studies that have delimited the particular activities where embodiment supports learning with technology. Thus, this paper reports on a study that asked people to physically enact with their bodies visual material they were studying. It compares learning from both static and dynamic visualizations as it has been argued that such an approach may be particularly useful with animations and video (de Koning & Tabbers, 2011). Moreover, by measuring the types of body actions learners make and relating these to the outcomes of learning, we hoped to explore not just whether enactment helps learning from visual materials but why it might (or might not) be helpful in this case.

1.1 Enactment as a learning strategy

It is now well established that learning from visual materials is more effective if students are encouraged to study them using an appropriate learning strategy (e.g., for overarching reviews

1
2
3
4 see Chi, 2009; Dunlosky, Rawson, Marsk, Nathan, & Willingham, 2013; Renkl & Scheiter,
5
6
7 2017). For example, students' learning is enhanced if they engage in inferential processes
8
9 brought about when they self-explain diagrams (e.g., Author C, 2003; Cromley et al., 2013).
10
11 Similarly, learning from animations is increased if students are taught (and then apply) strategies
12
13 that encourage them to select, organise, transform and integrate the information depicted
14
15 (Kombartzky, Ploetzner, Schlag, & Metz, 2010) or when they draw key aspects of the animation
16
17 (Mason, Lowe, & Tornatora, 2013). When learners do not apply such strategies, extensive
18
19 research suggests that far too frequently people can fail to attend to important information in the
20
21 visualizations, either skipping them altogether (e.g., Hegarty & Just, 1993; Schmidt-Weigand,
22
23 Kohnert, & Glowalla, 2010); attending to superficial salient details, not relevant information
24
25 (e.g., Hegarty, Canham, & Fabrikant, 2010; Lowe, 2004) or failing to integrate them with other
26
27 representations (e.g., Author C, 2006; Seufert, 2003). In this study of learning from complex
28
29 visual materials, we asked whether using your body in a meaningful way (enactment) could also
30
31 be as a successful learning strategy. We use the term 'enactment' as it encompasses many ways
32
33 of using one's body to express and development understanding rather than 'gesturing' as this is
34
35 often defined to include only hand and or occasionally head movements (e.g. "Gestures are
36
37 spontaneous hand movements that accompany speech", Novack & Goldin-Meadow, 2017, p.
38
39 652). The exception to this terminological approach is when we describe the research of others
40
41 and hence use the terms that authors themselves used.
42
43
44
45
46
47
48
49

50 Interest in embodied learning is increasing as research has found that the enacting¹ of
51
52 knowledge and concepts through bodily activity can support learning. For example, when people
53
54 gesture as they solve problems and reason, their understanding is enhanced (e.g., Goldin-
55
56 Meadow & Wagner, 2005; Stieff, Lira, & Scopelitis, 2016). Chu and Kita (2011) asked people to
57
58
59
60
61
62
63
64
65

1
2
3
4 solve 3d mental rotation tasks and found that those who spontaneously gestured solved difficult
5
6 problems more accurately. Moreover, this is unlikely to be epiphenomenal as, in a follow-up
7
8 study, when participants were encouraged to gesture, they solved these problems more accurately
9
10 than people who were not permitted to gesture. Furthermore, even when the participants were
11
12 prohibited from gesturing in a subsequent problem-solving phase this advantage was maintained.
13
14 These benefits of gesturing are not only found in laboratory tasks. Cook, Mitchell, and Goldin-
15
16 Meadow (2008) found that requiring children to gesture whilst they solved mathematics
17
18 problems involving grouping was associated with better performance four weeks later. Gesturing
19
20 as a learning strategy can also be trained. For example, Stieff et al. (2016) taught students
21
22 gestures, which they then practiced, that represented the spatial relationships depicted in
23
24 diagrams of molecular representations. These students performed better than those who studied
25
26 alone or who observed a teacher who gestured these relationships. They did equally well to those
27
28 taught with concrete models and better in an exam when concrete models were not available.
29
30
31
32
33
34
35

36 Research has also addressed whether enactment and gesturing can also help learning from
37
38 visual materials. This research can be classified in terms of how human movement is employed
39
40 in the studies. The majority of the research has explored whether the observation of included
41
42 gestures or representations of hands in learning materials facilitates students' understanding (e.g.,
43
44 de Koning & Tabbers, 2011; Pi, Hong, & Yang, 2017; Schwan & Riempp, 2004). This research
45
46 typically does find a benefit; although not invariably so (see Castro-Alonso, Ayres, & Paas, 2014
47
48 for a counterexample). However, we are interested in whether asking learners themselves to
49
50 move facilitates learning and here the research is decidedly mixed. Post, van Gog, Paas and
51
52 Zwaan (2013) compared young children learning about grammar from animations when they
53
54 either saw an animation or saw the animations which also included procedural information about
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 grammar rules in gestures which the children had to mimic. Contrary to their expectations, there
5
6 was no overall benefit of gesturing and indeed for children who began the study with lower
7
8 language skills this activity even harmed their performance on later tests. De Koning and
9
10 Tabbers (2013) compared participants learning from an animation of the lightening cycle when
11
12 they either had to watch the animation, watch the animation where important movements were
13
14 cued by a pointing with a photograph of a human hand or watch the animation where important
15
16 movements were cued by a pointing arrow that they also needed to gesture towards. Again,
17
18 gesturing along with the arrow was no more successful than simply watching the animation (and
19
20 not as successful as watching a human hand). The authors speculated that any potential benefits
21
22 of gesturing may have been obscured by the unusual nature of this constrained gesturing activity.
23
24
25
26
27

28
29 In contrast, Macken and Ginns (2014) do find that gesturing activities including tracing
30
31 over diagrams and between text and diagrams enhanced learning about the human heart. Similar
32
33 gesturing and tracing activities were used by Hu, Ginns and Bobis (2015) who also found an
34
35 advantage for these activities in the context of geometry worked-out examples. Moreover,
36
37 Lajevardi, Narang, Marcus, and Ayres (2017) do find a benefit of making gestures whilst
38
39 studying visual materials. In their studies, learners were studying how to write unfamiliar
40
41 characters (in Mandarin and Persian) and some additionally had to mimic the writing actions
42
43 they observed. However, half the participants observed animations whereas the others observed
44
45 static materials which showed the keyframes but no movements. Their results indicated that
46
47 gesturing was beneficial in the static but not animated conditions. Existing research is therefore
48
49 mixed, suggesting that although learners can benefit from enactment, there is only a limited
50
51 number of situations where this approach has been found to enhance learning from visual
52
53 materials.
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Although this may initially seem disappointing, it is perhaps not unexpected as research
5
6 has shown that even those learning strategies that have utility across different study materials,
7
8 such as self-explanation or drawing have been found to have boundary conditions and situations
9
10 where they do not help learning (e.g., Ploetzner & Fillisch, 2017). We suggest that one way of
11
12 clarifying this is to consider the mechanisms by which enactment may help learning; considering
13
14 both fairly general mechanisms as well as very specific ones.
15
16
17

18
19 General mechanisms by which enactment could support learning from visual materials
20
21 include increasing learners' motivation to study. Being asked to gesture and enact is, in our
22
23 experience, a rather unusual approach to support learning and one that students will rarely
24
25 encounter in formal school. Consequently, being encouraged to do so in experimental situations
26
27 may serve to enhance motivation in general due to its novel nature. Similarly general are the
28
29 benefits that might come from increasing participants' cognitive and physiological arousal in line
30
31 with predictions from arousal theory (e.g., Yerkes & Dodson, 1908; Lambourne &
32
33 Tomporowski, 2010). Potentially, increasing one's physical activity to greater levels might
34
35 enhance cognitive performance, especially as such activity is unlikely at maximal levels. These
36
37 types of benefit would be expected to accrue irrespective of how participants respond and the
38
39 nature of the learning materials.
40
41
42
43
44
45

46 More specific benefits of enactment, however, would depend upon the nature of learners'
47
48 enactment and how this aligns with the form and content of the studied material. One argument
49
50 for the benefits of gesturing to learn is that people are particularly attuned to learn about human
51
52 movement through bodily experiences, either their own or others (de Koning, & Tabbers, 2011).
53
54 This, it is suggested, explains why people are particularly good at learning by watching
55
56 animations of complex motor and procedural skills such as people tying knots (Schwan &
57
58
59
60
61
62
63
64
65

1
2
3
4 Riempp, 2004) or building with construction toys (Castro-Alonso, Ayres, & Paas, 2015). This
5
6 argument often recruits an explanation based upon activation of the human mirror neuron system
7
8 (e.g., Ayres, Marcus, Chan, & Qian, 2009). For example, Author B (2015) used functional Near-
9
10 Infrared Spectroscopy (fNIRS) and showed increased activation in the inferior-frontal cortex of
11
12 participants low in visual-spatial abilities when they watched instructional gestures. If human
13
14 movement is not shown but people need to enact something dynamic they have seen, it would be
15
16 predicted that the success of this approach would depend upon whether these movements are
17
18 ones they can successfully embody. Consequently, this predicts that the content of the studied
19
20 material will have a strong influence on the likelihood of enactment being a successful way to
21
22 learn.
23
24
25
26
27

28
29 Enactment may be particularly helpful when learning with visual materials and especially
30
31 animations as previous research points out these can both “underwhelm” or “overwhelm”
32
33 learners (Lowe, 2003). When visual materials underwhelm, learners pay insufficient attention to
34
35 them, focus on surface and superficial features and process them only shallowly (e.g., Harp &
36
37 Mayer, 1997; Sanchez & Wiley, 2006). This tendency is particularly marked with animations
38
39 (Lowe, 2003). Consequently, being required to enact studied visual materials with one’s own
40
41 body may focus learner’s attention on the important aspects of the visual materials, in similar
42
43 ways to other constructive strategies such as drawing (Wiley, 2019). Alternatively, animations,
44
45 are notorious for the way that their transient nature burdens learners’ limited working memory
46
47 capacity (Lowe, 2003). It is argued that using one’s body to represent the contents of animation
48
49 can offload cognition from limited working memory capacity onto one’s hands, hence freeing up
50
51 cognitive resources for more extensive processing of the animation (e.g., Lajevardi, Narang,
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Marcus, & Ayres, 2017; de Koning & Tabbers, 2011, 2013; Pouw, de Nooijer, van Gog, Zwaan,
5
6 & Paas, 2014).

7
8
9 A final level of explanation concerns what learners specifically do with their bodies: that
10
11 is, whether their gestures and enactments are representational and appropriate. Kita, Alibali and
12
13 Chu (2017, p. 245) suggest that gestures can help learning when they act as abstracted
14
15 representational motor actions that depict “action, motion, or shape, or that indicate location or
16
17 trajectory”. Similarly, Novack and Goldin-Meadow (2017) argue that as representational
18
19 gestures abstract away from the specifics of the situation to highlight its key properties this can
20
21 support understanding, retention, and generalization. In this way, they can become, as Pouw et al
22
23 (2014) argue, epistemic actions that reveal by their representational properties aspects of a
24
25 situation or phenomenon that might otherwise not be made visible. This explains why some
26
27 representational gestures are more effective than others (Goldin-Meadow, Cook, & Mitchell,
28
29 2009). These movements can be relatively simple such as the V-shaped gestures to components
30
31 of mathematics problems (e.g., Goldin-Meadow et al., 2009) or they can be more complex
32
33 roleplays such as when children enact the orbit of the moon around the earth or the movement of
34
35 water molecules in ice (Tytler, Prain, Hubber, & Waldrip, 2013). Moreover, this body-based
36
37 knowledge may result not only in an extra sensory code in memory (Engelkamp & Zimmer,
38
39 1984) but it may also result in new conceptual metaphors (e.g., Lakoff & Johnson, 1980;
40
41 McNeil, 1992) that help learners make sense of a situation in a new way. For this to be
42
43 successful though, these actions must be congruent with the core conceptual structures to be
44
45 learned (Lindgren & Johnson-Glenberg, 2013; Pouw et al, 2014).

46
47
48 In elucidating these arguments and presenting them as separate aspects of enactment we
49
50 do not mean to imply that bodily action serves only one purpose at a time; for example, learners
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 may be motivated by the physical activity involved in making new representational gestures that
5
6 embody key aspects of the studied material. Nonetheless, in designing our study we paid close
7
8 attention to these different mechanisms so that we could manipulate variables that may impact
9
10 upon them and measured both the processes and outcomes of enactment-based learning in an
11
12 attempt to distinguish whether enactment had more general or specific effects in this case.
13
14
15
16
17

18 19 **1.2 Hypotheses**

20
21 The material learners studied in this experiment concerned fish locomotion: that is,
22
23 participants studied visual representations of fish movements in order to identify different
24
25 patterns (for previous studies with same learning task and similar materials see Authors A & B,
26
27 2011, 2012, 2013, 2014). Recognizing such locomotion patterns is non-trivial as learners must
28
29 consider both temporal aspects and spatial aspects that form the basis of species classification
30
31 (Videler, 1993). Thus, the content of the studied material is about bodily movement but,
32
33 similarly to the classic example where people fail to understand horse locomotion, participants
34
35 are required to understand bodies very different from their own. Consequently, if enactment does
36
37 help learning the explanation cannot rest on those mechanisms proposed for understanding
38
39 human movement (e.g., de Koning & Tabbers, 2011) but instead still could help learning through
40
41 the other general and specific mechanisms we articulated above.
42
43
44
45
46
47

48 H1: Participants who enact fish locomotion in their own bodies will have enhanced
49
50 learning outcomes compared to those who do not.
51
52

53 We used different forms of visual material, not because we were interested in whether
54
55 dynamic or static study materials help people learn best as many other studies have addressed
56
57 that issue (e.g., Höffler & Leutner, 2007; Berney & Bétrancourt, 2016). We did so we could
58
59
60
61
62
63
64
65

1
2
3
4 specifically test whether dynamic and static materials benefited most from enactment. Given the
5
6
7 mixed results reviewed above, we formed two alternative hypotheses concerning how enactment
8
9 would influence learning from the different forms of visual materials.
10

11 H2a: Participants who watch videos whilst enacting will have enhanced learning
12
13 outcomes compared to those who view photographs whilst enacting.
14
15

16 We made this prediction on the basis that videos are transient representations. Hence
17
18 participants will benefit from offloading their cognition onto their hands (consequently we take
19
20 self-reported measures of mental effort) and in attempting to emulate the motion they see they
21
22 may place closer attention to the video and process it less superficially. Previous research which
23
24 found a benefit for gesturing from a video (Lajevardi et al., 2017) also indicated that this benefit
25
26 only manifested when the materials studied were at least moderately complex and therefore we
27
28 also expected that this may be true in our materials.
29
30
31

32
33 H2b: Participants who view photographs whilst enacting will have enhanced learning
34
35 outcomes compared to those who watch videos whilst enacting.
36
37

38 In this case, participants viewing photographs do not have access to the full dynamic
39
40 information provided by the videos and so by using their bodies as an active strategy, this could
41
42 compensate for this missing information.
43
44

45 In order to test whether the benefits of enactment were (in part at least) due to learners
46
47 forming schematic representations of the fish locomotion in their own bodies, we also videoed
48
49 the participants whilst they studied the material. The enactments they made were coded in terms
50
51 of their presence or absence and their congruency (see section 2.3). Consequently, we made the
52
53 following predictions.
54
55
56
57
58
59
60
61
62
63
64
65

H3: As video provides a fuller indication of the fish movements, participants in the video condition will make more congruent enactments than those viewing still photographs.

H4: Participants who make congruent enactments will learn more than those who make less congruent ones.

2. Method

2.1 Participants and design

Eighty-five students (with different majors from a German University) participated in the study for either payment (16 €) or course credit. Two participants had to be excluded from data analyses: One participant did not comply with the instruction and did not perform any gestures when being prompted to do so. The other participant was a Major in Biology and possessed too much background knowledge regarding the domain of the study. This left 83 participants (59 female; $M = 24.16$ years, $SD = 4.24$). Given that we were interested in how enactments would enhance learning from different visual formats (i.e., static versus dynamic materials), these two factors were varied experimentally in a 2x2 between-subjects design. That is, students learned from either photographs (static) or videos (dynamic), thereby varying the *visual format* of instruction. Furthermore, half of the students were prompted to perform enactments of what they observed in the visuals, while the other half received no enactment prompts (*enactment prompt*). The participants were randomly assigned to one of the four conditions. There were 22 participants in the condition without enactment instruction and photographs, 21 participants in the condition without enactment prompt and videos, and 20 participants in the two remaining conditions.

2.2 Materials

The study materials consisted of six videos, one per fish, whereby each fish deployed a different locomotion behavior (i.e., a pattern of moving forward). The movement patterns varied with respect to the body parts a particular fish uses to generate movement (i.e., whole body, pectoral fins or dorsal and anal fins) and the way in which it moves these body parts (i.e., in a wave-like, undulating or paddle-like, oscillating manner): (1) *anguilliform*: undulation of the body as a whole with constant amplitudes (e.g., eel); (2) *subcarangiform*: undulation of the body as a whole with decreasing amplitudes (e.g., trout); (3) *rajiform*: undulation of enlarged pectoral fins, mostly in addition with oscillating movements of these fins (e.g., ray); (4) *balisitifform*: undulation of the dorsal and anal fins (e.g., triggerfish); 5. *tetraodontiform*: oscillation of the dorsal and the anal fins, as well as sometimes undulation of the pectoral fins (e.g., puffer fish); and 6. *labriform*: oscillation of the pectoral fins (e.g., wrasse). One of the major challenges in identifying these locomotion patterns is that fish may deploy a variety of additional movements, for instance, in order to navigate, which may easily be confused with these characteristic movements pertaining to propulsion.

Depending on the condition, there was either a single video of each fish or a series of sequentially presented photographs. The videos were non-interactive underwater videos showing a fish performing one complete movement cycle. The photographs consisted of stills extracted from the videos (see Figure 1 for an example of the *tetraodontiform* movement pattern), which corresponded to the key states of the movement as determined by two domain experts (cf. Authors A & B, 2012). For each movement pattern, nine stills were chosen. Each still was presented for three seconds and then automatically replaced by the next picture in the sequence.

1
2
3
4 For learning, the visualizations of all six fish were shown once (trial 1) and then a second time
5
6
7 (trial 2; see procedure for details).
8
9

10 11 **2.3 Measures**

12
13
14 As control variables, we assessed participants' spatial abilities, their verbal abilities, and
15
16 their familiarity with related tasks and contents. Moreover, we asked the participants whether
17
18 they had received training in either dancing or acting since this might influence their propensity
19
20 to use their body as a way of expressing themselves. There were no differences between
21
22 conditions regarding the number of participants who had received either type of training. As
23
24 dependent variables, we assessed learning outcomes by means of a movement pattern
25
26 classification task (recognition) with a range of difficulty and two open-ended near and far
27
28 transfer written description tasks. Furthermore, self-reported mental effort during learning was
29
30 registered. Finally, the body movements that students produced were coded. Each of these is
31
32 explained in more detail below.
33
34
35
36
37

38 **Spatial ability.** Participants' spatial ability was assessed with a shortened paper-pencil
39
40 version of the Paper Folding Test (PFT, Ekstrom, French, Harman, & Dermen, 1976). It
41
42 consisted of 10 multiple-choice items. Each item showed a square piece of paper being folded
43
44 and then punched with a pencil as well as five possible variants of that same paper when the
45
46 paper is completely unfolded. Participants were asked to determine which of the five versions
47
48 correctly represented the folded paper stimulus. They had three minutes to work on the PFT. For
49
50 each correct answer, a participant received one point and for each incorrect answer, one point
51
52 was deducted, yielding a minimum of -10 points and a maximum of +10 points.
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Verbal abilities.** Participants' language skills were tested with the *Mehrfachwahl-*
5 *Wortschatz-Intelligenz-Test* (MWT-B, Lehl, 2005). The MWT-B consists of 37 items, whereby
6
7 for each item there are four non-words (meaningless letter combinations) and one meaningful
8
9 word. The task was to identify the meaningful word. There were no time limits for this test. For
10
11 each correct answer, a participant received one point (max. 37 points).
12
13

14
15
16 **Familiarity with the domain.** The questionnaire to assess participants' familiarity with
17
18 the domain assessed the number of courses taken in Biology and Physics (details of participants'
19
20 school education), experiences with the marine world (e.g., diving, snorkeling, swimming,
21
22 rowing, and/or owning an aquarium), interest in related topics (e.g., biology, zoology, physics,
23
24 aircraft construction, and/or shipbuilding) and media use (e.g., documentaries and/or books on
25
26 fish or sea life). Participants could achieve a total amount of 25 points in this questionnaire,
27
28 which served only to identify students in the unlikely event that they had expertise in this
29
30 otherwise rather exotic domain.
31
32
33

34
35 **Learning outcomes.** Recognition was assessed in a movement pattern classification test.
36
37 In this test, participants saw 28 underwater videos of fish that displayed the same movements as
38
39 the previously studied fish but that differed their visual appearance (i.e., shape and color). There
40
41 were four to six videos for each to-be-learned movement pattern. Students had to classify the
42
43 movement pattern of each fish by assigning it to the corresponding movement pattern from the
44
45 learning phase. This was done non-verbally: for each movement pattern from the learning phase,
46
47 one still was extracted from the videos to represent the respective movement pattern. Thus,
48
49 students did not have to be able to name the movement patterns; rather, they could classify a test
50
51 item simply by assigning it to the representative photograph.
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 The 28 items were chosen so that they represented different levels of task difficulty. In
5
6 easy items (13 items) the depicted fish showed the locomotion pattern for propulsion
7
8 continuously, and the fish showed no other movements. Data for four of these items showed too
9
10 little variance as almost all students recognized the fish correctly (i.e., ceiling effect); therefore,
11
12 these items had to be excluded from the analysis. In the medium difficulty items (7 items), the
13
14 depicted fish showed the relevant locomotion pattern continuously, but they also performed
15
16 some movements that were similar to another locomotion pattern. In difficult items (8 items) the
17
18 depicted fish either showed the pattern relevant for propulsion continuously, but contained
19
20 additional movements similar to at least two other locomotion patterns, or the depicted fish did
21
22 not show the relevant locomotion pattern for propulsion continuously or did not show it in a
23
24 salient manner. Participants received one point for each correct answer, yielding 9, 7, and 8
25
26 points for easy, medium, and difficult items, respectively, which were transformed into ratios for
27
28 easier interpretation.
29
30
31
32
33
34
35

36 Additionally, learning outcomes in terms of near and far transfer were assessed with
37
38 written open description tasks. In the near transfer test, participants were shown videos of four
39
40 novel fish that performed the same continuous propulsions as the fish seen during learning. In the
41
42 far transfer test, participants were shown videos of four marine animals (three fishes and a cuttle)
43
44 that showed different movement behaviors to the ones from the learning phase. In both tasks,
45
46 participants were asked to describe the fish's movements as precisely as possible and received
47
48 points whenever they recognized which body part is important for generating propulsion and
49
50 which body part is important for navigation (1 point for each item). Additionally, they received
51
52 points whenever they described correctly how these body parts moved. For 3 out of 4 items in
53
54 each test, this description required mentioning one aspect (1 point for each item), whereas in one
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 item it required mentioning two aspects (2 points). Thus, there was a maximum score of 9 for
5
6 both transfer tests. The total score was converted into a ratio for easier interpretation.
7
8

9 **Mental effort.** Mental effort was assessed during learning separately for each of the six to-
10
11 be-learned locomotion patterns. Thus, participants were asked to rate how much effort they had
12
13 invested during learning on a 7-point Likert scale from “none at all” to “very much” after each
14
15 movement pattern was presented during the first learning trial.
16
17

18 **Enactment coding.** Participants were videotaped during the whole experiment with two
19
20 cameras positioned at different viewing angles so that their body movements during learning
21
22 could be coded. First, one of the authors of this paper determined for each of the 6 movement
23
24 patterns shown during learning whether students had performed enactments and if so, whether
25
26 they appeared related to the fish movement (relevant) or unrelated (e.g., communicative or other
27
28 gestures). The number of fishes for which students performed a movement related to the
29
30 locomotion of the fish at least once (that is, either during trial 1 or 2 or during both trials) was
31
32 used as a dependent measure. In cases where a participant displayed more than one enactment,
33
34 the two coders determined which was the dominant one and would thus be chosen for further
35
36 analysis. In cases of disagreement, a consensus was achieved by discussion. For the chosen
37
38 enactments, coders decided if the enactment was congruent with the fish movement (0 = not at
39
40 all, 1 = some, 2 = close match).
41
42
43
44
45
46
47
48
49

50 **2.4 Procedure**

51
52 Participants were tested in single sessions. To allow them to use their whole body, while at the
53
54 same giving them the opportunity to sit during the experiment, the following set-up was chosen.
55
56

57
58 The computer on which the materials were displayed was placed on a table whose height was
59
60
61
62
63
64
65

1
2
3
4 adjusted so that a person could easily stand in front of it while watching the screen. Placed in
5
6 front of the table was a bar stool that allowed maximum mobility with the upper body part whilst
7
8 sitting on it. Also, students could easily move it away if they preferred to gesture while standing.
9
10
11 Two video cameras were positioned to the left and at the right side of the table, recording the
12
13 participants from the front and each side.
14
15

16
17 Once students were comfortable with this arrangement, they were first given a short
18
19 computer-based written overview on the experiment and an instruction that they were to learn as
20
21 much as possible about the movement patterns deployed by different fish so that they later would
22
23 be able to recognize and describe these patterns as accurately as possible. Participants were
24
25 furthermore informed that they could watch the videos/photographs sequences for as long as they
26
27 liked and that each video/photograph sequence would restart automatically until they pressed the
28
29 space to continue to the next movement pattern. In addition, participants in the conditions with
30
31 enactment prompt received the following instruction: “Whilst you watch the visualizations,
32
33 please use your hands and body to gesture in a way that you think will help you to better
34
35 understand the fish’s movements. This might sound strange; however, previous research has
36
37 shown that gestures deployed in such a way may improve learning and that they do so
38
39 particularly if they are matched to what is to-be-learned”¹.
40
41
42
43
44
45

46
47 The instructions were followed by the learning phase, in which the visualizations for the
48
49 six movement patterns were presented at the center of the computer screen. After each
50
51 presentation of the six movement patterns, participants were asked to rate their mental effort
52
53 concerning it. These sequences of six movements patterns were then presented for a second time.
54
55 After the learning phase, participants first performed the near transfer task. They were instructed
56
57 to describe the movements of the fish shown in the four videos as precisely as possible.
58
59
60
61
62
63
64
65

1
2
3
4 Afterward, participants performed the classification task. They were instructed to classify the
5
6 movement patterns depicted in this task by comparing them with the movement patterns of the
7
8 learning phase. Then, participants performed the far transfer task. They were again instructed to
9
10 describe the movements of the fish shown in the four videos as precisely as possible. No time
11
12 restrictions were giving for any of the learning outcome tests. Finally, participants answered the
13
14 PFT and the MWT-B. Then they were debriefed and thanked for their participation. A single
15
16 session took approximately 90 minutes.
17
18
19
20
21
22

23 24 **3. Results**

25
26 The first step in the analysis was to check whether the conditions differed regarding the students'
27
28 verbal and spatial abilities and their domain familiarity (Table 1). A multivariate ANOVA with
29
30 visual format (photographs vs. video) and enactment prompt (with vs. without) as factors
31
32 revealed no effects of either factor and no interaction (all $F_s < 1$). Correlational analyses between
33
34 students' entry characteristics and overall performance in the description task and the
35
36 classification task were performed. There were no significant correlations between the students'
37
38 verbal ability score and performance (description task: $r = .04$; $p = .74$; classification task: $r =$
39
40 $.005$; $p = .96$) or between domain familiarity and performance (description task: $r = .05$; $p = .65$;
41
42 classification task: $r = -.17$; $p = .13$). In line with the literature (cf. Höffler, 2010), spatial ability
43
44 as measured by the PFT, was positively correlated with performance (description task: $r = .32$; p
45
46 $= .004$; classification task: $r = .37$; $p = .001$), which is why it was considered as a covariate² in
47
48 the subsequent analyses.
49
50
51
52

53 54 **3.1 Mental effort, enactment quantity and congruency**

55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Students' averaged ratings regarding their experienced mental effort were analyzed with
5
6 a two-factorial ANCOVA with visual format and enactment prompt and spatial ability as
7
8 covariate. Means and standard deviations are displayed in Table 2. There were no significant
9
10 effects (enactment prompt: $F(1,78) = 1.33$, $MSE = 1.82$, $p = .25$, $\eta^2 = .02$; visual format: $F(1,78)$
11
12 $= 3.06$, $MSE = 1.82$, $p = .08$, $\eta^2 = .04$; enactment prompt \times visual format: $F < 1$; spatial ability:
13
14 $F(1,78) = 2.86$, $MSE = 1.82$, $p = .095$, $\eta^2 = .04$). Thus, neither the format of the studied visual
15
16 material or the prompt to enact seemed to affect participants' experiences of effort.
17
18

19
20
21 Furthermore, the number of fishes where participants performed a relevant body
22
23 movement at least once was determined (Table 2). In the prompted conditions, this coding was
24
25 used to identify the single person who had not complied with the prompt, who was then excluded
26
27 from all analyses. However, in the conditions without the prompt, 20 out of 43 participants
28
29 performed at least one enactment despite not being told to do so. We decided against removing
30
31 these participants as this would not only have drastically reduced our sample size, it would also
32
33 have yielded a non-representative sample.
34
35
36
37

38
39 The number of fishes with enactment was analyzed using a two-factorial ANCOVA with
40
41 visual format and enactment prompt as factors, and spatial ability as covariate. As expected, the
42
43 prompt increased the number of enactments (with prompt: $M = 5.92$, $SE = 0.62$; without prompt:
44
45 $M = 2.15$, $SE = 0.30$), $F(1,78) = 74.37$, $MSE = 3.95$, $p < .001$, $\eta^2 = .49$. Hence, our experimental
46
47 manipulation had been successful in that those students who had been prompted to enact the
48
49 observed movements performed enactments more frequently. Visual format had no significant
50
51 effect on the number of enactments performed nor was there an interaction between visual
52
53 format and enactment prompt (both $F_s < 1$). Also spatial ability had no effect on the number of
54
55 enactments $F(1,78) = 1.57$, $MSE = 3.95$, $p = .21$, $\eta^2 = .02$.
56
57
58
59
60
61
62
63
64
65

1
2
3
4 In the next step, we analyzed the mean correspondence between the participants
5
6 movements and the actual movement displayed by the fish as a measure of congruency using the
7
8 same analysis as above. This variable was of interest, because we had assumed that videos would
9
10 yield more congruous enactments than stills (Hypothesis 3) and that students performing more
11
12 congruous enactments would also show better learning (cf. Hypothesis 4). Note that in the
13
14 conditions without prompts, there were only 20 students who performed spontaneous
15
16 enactments; thus, the sample size in the not prompted conditions is reduced to $n = 10$ for each
17
18 visual format condition. There were no main effects of spatial ability, $F(1,55) = 1.16$, $MSE =$
19
20 0.97 , $p = .29$, $\eta^2 = .02$, visual format ($F < 1$), or enactment prompt, $F(1,55) = 1.12$, $MSE = 0.97$, p
21
22 $= .30$, $\eta^2 = .02$. However, there was a significant interaction between visual format and enactment
23
24 prompt, $F(1,55) = 6.05$, $MSE = 0.97$, $p = .02$, $\eta^2 = .10$. Bonferroni-adjusted posthoc comparisons
25
26 revealed that in the conditions without prompts, visual format had no effect on congruency
27
28 (photographs: $M = 1.31$, $SE = 0.10$; video: $M = 1.17$, $SE = 0.10$; $p = .34$). In prompted conditions,
29
30 however, the enactments were more congruent when based on videos ($M = 1.47$, $SE = 0.07$) than
31
32 when based on sequences of photographs ($M = 1.19$, $SE = 0.07$; $p = .006$). Hence, in line with
33
34 Hypothesis 3, videos yielded a more comprehensive mental representation of the observed
35
36 movements that was expressed in the congruency of the enactments performed by students.
37
38
39
40
41
42
43
44

45
46 To summarize, even though those students who were prompted to enact while watching
47
48 the visualizations did so to a greater extent than those who were not instructed, in the latter
49
50 condition almost 50% of students did so nevertheless. There were no differences between
51
52 spontaneous and prompted enactments regarding their quality; however, enactments
53
54 corresponded best with the fish movements when students were prompted to perform them based
55
56 on videos. Thus, if congruency of enactments mattered for learning (as assumed in Hypothesis
57
58
59
60
61
62
63
64
65

1
2
3
4), then we would expect students in the video (compared to the photograph) conditions to show
5
6
7 better learning outcomes when performing enactments.
8

9 **3.2 Learning outcomes**

10
11 Learning outcome measures referred to the students' performance in the movement
12
13 classification task and in the description task (Table 3).
14
15

16 The students' performance in the movement classification task was analyzed using a two-
17
18 factorial repeated-measures ANCOVA with visual format and enactment prompt as factors, task
19
20 difficulty (low, medium, high) representing the repeated measure and spatial ability as the
21
22 covariate.
23
24

25
26 There was a main effect of task difficulty, $F(2,156) = 51.39$, $MSE = 0.02$, $p < .001$, $\eta^2 =$
27
28 $.40$, suggesting that tasks classified as easy were indeed solved at a much higher rate ($M = .93$,
29
30 $SE = .01$) than tasks classified as either being of medium ($M = .68$, $SE = .02$, $p < .001$) or high
31
32 difficulty ($M = .68$, $SE = .02$, $p < .001$), with the latter not differing from each other³. There were
33
34 no main effects of visual format, $F < 1$, or enactment prompt, $F(1,78) = 2.87$, $MSE = 0.05$, $p =$
35
36 $.09$, $\eta^2 = .04$ nor was there an interaction among the two factors, $F < 1$. However, task difficulty
37
38 interacted with prompt condition, $F(2,156) = 3.27$, $MSE = 0.02$, $p = .04$, $\eta^2 = .04$. The
39
40 Bonferroni-adjusted posthoc analyses revealed that there were no effects of being prompted to
41
42 enact on easy tasks (without prompts: $M = .93$, $SE = .02$, with prompts: $M = .93$, $SE = .02$) or
43
44 difficult tasks (without prompts: $M = .66$, $SE = .03$, with prompts: $M = .70$, $SE = .03$). However,
45
46 for classification tasks of medium difficulty, students performed better after having received
47
48 enactment prompts ($M = .73$, $SE = .03$) compared with the no-prompt conditions ($M = .63$, $SE =$
49
50 $.03$, $p = .03$). There was no interaction between visual format and task difficulty, $F(1,78) = 2.51$,
51
52 $MSE = 0.02$, $p = .08$, $\eta^2 = .03$, and no three-way interaction, $F < 1$. Finally, spatial ability had a
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 strong positive impact on the performance in the classification task, $F(1,78) = 12.53$, $MSE =$
5
6 0.05 , $p = .001$, $\eta^2 = .14$. Accordingly, as expected in Hypothesis 1, being prompted to perform
7
8 enactments did improve learning outcomes as assessed with medium-difficulty classification
9
10 tasks. In contrast to the two conflicting Hypotheses 2a and 2b, however, these effects of
11
12 enactment were independent of the visual format.
13
14

15
16 Performance in the description task was analyzed with a two-factorial ANCOVA with
17
18 visual format and enactment prompt as factors, and spatial ability as the covariate. It showed no
19
20 effects of enactment prompt and no interaction of enactment prompt and visual format, both $F_s <$
21
22 1. However, performance was affected by visual format, $F(1,78) = 6.63$, $MSE = 0.03$, $p = .01$, η^2
23
24 $= .08$, in that learners with a sequence of photographs ($M = 0.70$, $SE = 0.03$) performed better
25
26 than those learning with videos ($M = 0.61$, $SE = 0.03$). Again, spatial ability had a strong positive
27
28 influence on the performance in the description task, $F(1,78) = 8.66$, $MSE = 0.03$, $p = .004$, $\eta^2 =$
29
30 $.10$. Accordingly, positive effects of enactment were limited to classification tasks.
31
32
33

3.3 Relationships between enactments and learning outcomes

34
35
36 In a final step, we analyzed the relationship between quantity and quality of the
37
38 participants' body movements and their learning outcomes on the medium difficulty
39
40 classification task. This was done by performing a regression analysis within the conditions with
41
42 the enactment prompts using the number of performed enactments and their correspondence with
43
44 the fish movement as predictors. The overall regression model was not significant, $F < 1$. Thus,
45
46 the present results do not confirm Hypothesis 4 according to which more congruent enactments
47
48 had been expected to yield better learning outcomes.
49
50
51
52
53
54
55
56
57

4. Summary and Discussion

58
59
60
61
62
63
64
65

1
2
3
4 This study was developed to explore whether enactment supported learning about non-
5
6 human movement from different forms of visual materials and if it did so, to attempt to establish
7
8 which of the many different explanations for enactment might be implicated.
9
10

11 We predicted being prompted to enact would increase learning outcomes (H1) and this
12
13 hypothesis did receive partial support from the findings as classification outcomes on medium
14
15 difficulty tasks were enhanced if learners enacted. However, this was the only measure to show
16
17 such a benefit as classification was not improved for easy or difficult classification items and the
18
19 description task. This finding is similar to that reported by Lajevardi et al. (2017) who also found
20
21 that tasks with easy movements did not benefit from enactment. It also suggests that enactment
22
23 (at least in this domain) will not result in richer verbal explanations as there was no effect of
24
25 enactment on the description task. This suggests only limited benefits for enactment. Notably,
26
27 these findings on enactment – where one uses the whole body to perform observed movements –
28
29 do not necessarily contradict earlier findings regarding the benefits of gestures (limited to hands
30
31 and heads) or tracing when learning from visual materials (e.g., Macken & Ginns, 2014; Hu et
32
33 al., 2015; Lajevardi et al., 2017). Rather, they suggest that benefits of using one’s body may be
34
35 strongly dependent on the type of movements performed relative to the task that needs to be
36
37 accomplished. For instance, tracing geometric figures is a tightly controlled activity with the
38
39 fingers and hands where the movement is constrained by the presented figure. In contrast, in the
40
41 present study students were free to choose how to enact the fishes’ movement using single parts
42
43 (e.g., their hands) up to their whole body. Future studies hence need to address the task
44
45 specificity of beneficial effects of using one’s body for learning.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Moreover, the interpretation of enactment effects in the present study is complicated by
5
6
7 the fact that whilst we had expected participants to be able to enact when prompted (which all
8
9
10 but one did), we had not anticipated so many spontaneous enactments; nearly half of the non-
11
12 prompted participants performed at least one such body motion, which was not influenced by
13
14 whether they saw static or dynamic materials. This was not expected as earlier studies with the
15
16 same materials had not reported seeing spontaneous enactment. Given the wide population
17
18 sampled from we do not propose the explanation here is program bleed, instead, we speculate
19
20 that the one key difference between this study and earlier ones with the same material was that
21
22 the participants were not seated in front of a desk in an office style chair. And their enactments
23
24 often involved using their bodies in ways that would not be easy in such a seated position.
25
26
27 Moreover, this suggests that many participants intuitively felt they would benefit from enactment
28
29 in this visually complex learning situation. This argument is support by Eielts et al (2018) who
30
31 found that spontaneous gesturing in increased when tasks demands increased. Consequently, it
32
33 seems probable that this spontaneous enactment served to reduce the predicted difference in
34
35 learning outcomes between the conditions. Therefore, a follow-up study could implement an
36
37 additional condition where enactment is not permitted as seen in studies such as Chu and Kita
38
39 (2011) and Eielts et al (2018) although this will now add an additional load of inhibiting a
40
41 presumably valued way of interacting with the material.
42
43
44
45
46
47

48 Earlier work had found mixed results concerning whether learning with videos or
49
50 photographs would most benefit from enactment and consequently, we could not made
51
52 alternative predictions for our second hypothesis (Hypothesis 2a and 2b; de Koning & Tabbers,
53
54 2013; Post et al., 2013; Lajevardi et al., 2017). We found no interaction between the form of the
55
56 materials and the enactment condition on learning outcomes suggesting that the benefits of
57
58
59
60
61
62
63
64
65

1
2
3
4 enactment on classification tasks were equally experienced with both visual materials. The
5
6
7 process measures we analyzed to try and delimitate different mechanisms for enactment effects
8
9
10 on learning provide further insight into these results.

11
12 One explanation that had previously been proposed for the benefits of enactment on
13
14 learning from animation, is that it reduces the cognitive demands on processing the animation
15
16 (e.g., Lajevardi et al., 2017; Pouw et al, 2014); however, participants' mental effort scores were
17
18 not affected by enactment prompts. Consequently, we think this explanation can be ruled out for
19
20 the present materials. A second explanation considered was that enactment would compensate
21
22 for the missing information in photographs as participants would generate this information in
23
24 their own motor movements. Again there does not seem to be evidence for this account either as
25
26 there are no differences in learning outcomes or mental effort between people who were
27
28 prompted to enact whilst viewing photographs and those who did not. There also was no
29
30 difference in spontaneous enactment by condition. However, we did find as predicted by H3 that
31
32 participants produced more congruent enactments (i.e., they produced body movements that
33
34 more closely resembled the fish movements that they observed) when they studied video
35
36 materials rather than static pictures. However, we did not find that congruency predicted learning
37
38 outcomes and so cannot find support for H4 that congruent enactments will be most beneficial. It
39
40 may be that fish bodies and human bodies are sufficiently dissimilar that schematically
41
42 representing the movement you observe (or infer) is not helpful and does not activate the human
43
44 mirror neuron system (e.g., Ayres et al., 2009; Brucker et al., 2015). Consequently, the (limited)
45
46 benefits of enactment for these materials appear most plausibly to come from fairly general
47
48 explanations of the advantages of enactment – that it is a motivating way to engage in studying
49
50 materials and that it may increase cognitive arousal. Additionally, participants may have studied
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 the materials more carefully if they needed to enact them. This material must be as neither too
5
6
7 easy (so that this additional work is not necessary) nor too difficult (so that learners do not
8
9 benefit from it) to be recognized. Subsequent studies could adopt additional process measures
10
11 (such as motivational questionnaires or eye tracking) to more directly test this explanation or
12
13
14 imaging techniques such as fNIRS.

15
16 A limitation of the present study pertains to the relatively low number of participants,
17
18 which can be argued to occlude possible effects of enactments. However, similar studies in the
19
20 field have used comparable sample sizes (e.g., Macken & Ginns, 2014, had 42 students randomly
21
22 assigned to two experimental conditions). Nevertheless, future studies should attempt to replicate
23
24 the present findings using larger samples to generate more robust findings.
25
26

27
28 In conclusion, this study has demonstrated that there is only limited benefit from enacting
29
30 non-human movements in participants' own bodies. At this point, there are still many
31
32 unanswered questions concerning when and how enactment supports visual learning and
33
34 additional studies are still required. However, we hope that we have responded to Lindgren and
35
36 Johnson-Glenberg's (2013) call for studies that look beyond the effects of an embodied approach
37
38 to technologically mediated learning to trying to understand what the mechanisms underlying
39
40 any effect might be and finding when embodiment may also not help learning.
41
42
43
44
45
46
47

48 **References**

- 49 Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning:
50 Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247-
51 286. <https://doi.org/10.1080/10508406.2011.611446>
52
53
54 Authors A & B 2011
55
56 Authors A & B 2012
57
58 Authors A & B 2013
59
60
61
62
63
64
65

1
2
3
4 Authors A & B 2014

5
6 Author B 2015

7
8 Author C 2003

9
10 Author C 2006

11 Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When
12 instructional animations are superior to equivalent static representations. *Computers in*
13 *Human Behavior*, 25, 348-353. <https://doi.org/10.1016/j.chb.2008.12.013>

14
15
16
17 Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-
18 analysis. *Computers & Education*, 101, 150-167.
19 <https://doi.org/10.1016/j.compedu.2016.06.005>

20
21
22 Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014). Learning from observing hands in static and
23 animated versions of non-manipulative tasks. *Learning and Instruction*, 34, 11–21.
24 <https://doi.org/10.1016/j.learninstruc.2014.07.005>

25
26
27 Castro-Alonso, J. C., Ayres, P., & Paas, F. (2015). Animations showing Lego manipulative tasks:
28 Three potential moderators of effectiveness. *Computers & Education*, 85, 1-13.
29 <https://doi.org/10.1016/j.compedu.2014.12.022>

30
31
32
33 Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating
34 learning activities. *Topics in Cognitive Science*, 1, 73-105. [https://doi.org/10.1111/j.1756-](https://doi.org/10.1111/j.1756-8765.2008.01005.x)
35 [8765.2008.01005.x](https://doi.org/10.1111/j.1756-8765.2008.01005.x)

36
37
38
39 Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving.
40 *Journal of Experimental Psychology. General*, 140, 102-116.
41 <https://doi.org/10.1037/a0021790>

42
43
44 Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last.
45 *Cognition*, 106, 1047–1058. <https://doi.org/10.1016/j.cognition.2007.04.010>

46
47
48 Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., &
49 Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram
50 comprehension skills: The critical role of inference while learning. *Learning and*
51 *Instruction*, 26, 45–58. <https://doi.org/10.1016/j.learninstruc.2013.01.003>

52
53
54
55 De Koning, B., & Tabbers, H. (2011). Facilitating understanding of movements in dynamic
56 visualizations: An embodied perspective. *Educational Psychology Review*, 23, 501-521.
57 <https://doi.org/10.1007/s10648-011-9173-8>

58
59
60
61
62
63
64
65

- 1
2
3
4 De Koning, B., & Tabbers, H. (2013). Gestures in instructional animations: A helping hand to
5 understanding non-human movements? *Applied Cognitive Psychology*, *27*, 683–689.
6 <https://doi.org/10.1002/acp.2937>
7
8
9
10 Dunlosky, J., Rawson, K. A., Marsk, E. A., Nathan, M. J., & Willingham, D. T. (2013).
11 Improving students' learning with effective learning techniques: Promising directions from
12 cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*, 4-
13 58. <https://doi.org/10.1177/1529100612453266>
14
15
16
17 Eielts, C., Pouw, W., Ouwehand, K., van Gog, T., Zwaan, R. A., & Paas, F. (2018). Co-thought
18 gesturing supports more complex problem solving in subjects with lower visual working-
19 memory capacity. *Psychological Research*. doi:10.1007/s00426-018-1065-9
20
21
22 Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-*
23 *referenced cognitive tests*: Princeton: Educational Testing Service.
24
25
26 Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory
27 unit. *Psychological Research*, *46*, 283-299.
28
29
30 Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley*
31 *Interdisciplinary Reviews: Cognitive Science*, *1*, 586-596. <https://doi.org/10.1002/wcs.55>
32
33
34 Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive*
35 *Sciences*, *9*, 234-241. <https://doi.org/10.1016/j.tics.2005.03.006>
36
37
38 Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas
39 about math. *Psychological Science*, *20*, 267-272. <https://doi.org/10.1111/j.1467->
40 [9280.2009.02297.x](https://doi.org/10.1111/j.1467-9280.2009.02297.x)
41
42
43 Harp, S. W., & Mayer, R. E. (1997). The role of interest in learning from scientific text and
44 illustrations: On the distinction between emotional interest and cognitive interest. *Journal of*
45 *Educational Psychology*, *89*, 92–102.
46
47
48 Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and
49 diagrams. *Journal of Memory and Language*, *32*, 717–742.
50
51
52 Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: how display
53 salience and knowledge affect performance in a graphic inference task. *Journal of*
54 *Experimental Psychology. Learning, Memory, and Cognition*, *36*, 37–53.
55
56
57 <https://doi.org/10.1037/a0017683>
58
59
60
61
62
63
64
65

- 1
2
3
4 Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-
5 analytic review. *Educational Psychological Review*, 22, 245-269.
6 <https://dx.doi.org/10.1007/s10648-010-9126-7>
7
8
9
10 Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-
11 analysis. *Learning and Instruction*, 17, 722–738.
12 <https://doi.org/10.1016/j.learninstruc.2007.09.013>
13
14
15 Hu, F.-T., Ginns, P., & Bobis, J. (2015). Getting the point: Tracing worked examples enhances
16 learning. *Learning and Instruction*, 35, 85-93.
17 <https://doi.org/10.1016/j.learninstruc.2014.10.002>
18
19
20
21 Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking?
22 The gesture-for-conceptualization hypothesis. *Psychological Review*, 124, 245-266.
23 <https://dx.doi.org/10.1037/rev0000059>
24
25
26
27 Kombartzky, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a
28 strategy for learning from animations. *Learning and Instruction*, 20, 424–433.
29 <https://doi.org/10.1016/j.learninstruc.2009.05.002>
30
31
32 Lajevardi, N., Narang, N. S., Marcus, N., & Ayres, P. (2017). Can mimicking gestures facilitate
33 learning from instructional animations and static graphics? *Computers & Education*, 110,
34 54-74. <https://doi.org/10.1016/j.compedu.2017.03.010>
35
36
37
38 Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual
39 system. *Cognitive Science*, 4, 195-208.
40
41
42 Lambourne, K., & Tomporowski, P. (2010). The effect of exercise-induced arousal on cognitive
43 task performance: a meta-regression analysis. *Brain Research*, 1341, 12-24.
44 <https://doi.org/10.1016/j.brainres.2010.03.091>
45
46
47
48 Lehl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B* (Vol. 5). Göttingen: Hogrefe.
49
50
51 Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for
52 research on embodied learning and mixed reality. *Educational Researcher*, 42, 445-452.
53 <https://doi.org/10.3102/0013189X13511661>
54
55
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4 Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and*
5 *Instruction, 14*, 257–274. <https://doi.org/10.1016/j.learninstruc.2004.06.003>
6
7
8 Macken, L., & Ginns, P. (2014). Pointing and tracing gestures may enhance anatomy and
9 physiology learning. *Medical Teacher, 36*, 596-601. doi:10.3109/0142159X.2014.899684
10
11 Mason, L., Lowe, R., & Tornatora, M. C. (2013). Self-generated drawings for supporting
12 comprehension of a complex animation. *Contemporary Educational Psychology*.
13 <https://doi.org/10.1016/j.cedpsych.2013.04.001>
14
15
16
17 McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of
18 Chicago Press.
19
20
21 Novack, M. A., & Goldin-Meadow, S. (2017). Gesture as representational action: A paper about
22 function. *Psychonomic Bulletin and Review, 24*, 652-665. [https://doi.org/10.3758/s13423-](https://doi.org/10.3758/s13423-016-1145-z)
23 [016-1145-z](https://doi.org/10.3758/s13423-016-1145-z)
24
25
26 Pi, Z., Hong, J., & Yang, J. (2017), Effects of the instructor's pointing gestures on learning
27 performance in video lectures. *British Journal of Educational Technology, 48*, 1020-1029.
28 <https://doi.org/10.1111/bjet.12471>
29
30
31 Ploetzner, R., & Fillisch, B. (2017). Not the silver bullet: Learner-generated drawings make it
32 difficult to understand broader spatiotemporal structures in complex animations. *Learning*
33 *and Instruction, 47*, 13–24. <https://doi.org/10.1016/j.learninstruc.2016.10.002>
34
35
36
37 Post, L. S., van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of simultaneously observing
38 and making gestures while studying grammar animations on cognitive load and learning.
39 *Computers in Human Behavior, 29*, 1450-1455. <https://doi.org/10.1016/j.chb.2013.01.005>
40
41
42
43 Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a
44 more embedded/extended perspective on the cognitive function of gestures. *Frontiers in*
45 *Psychology, 5*:359. doi:10.3389/fpsyg.2014.00359
46
47
48 Renkl, A., & Scheiter, K. (2017). Studying visual displays: How to instructionally support
49 earning. *Educational Psychology Review, 29*, 599–621. [https://doi.org/10.1007/s10648-015-](https://doi.org/10.1007/s10648-015-9340-4)
50 [9340-4](https://doi.org/10.1007/s10648-015-9340-4)
51
52
53
54 Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of
55 working memory capacity. *Memory and Cognition, 34*, 344–355.
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4 Schmidt- Weigand, F., Kohnert, A., & Glowalla, U. (2010). Explaining the modality and
5 contiguity effects: New insights from investigating students' viewing behaviour. *Applied*
6 *Cognitive Psychology*, 24, 226-237. <https://doi.org/10.1002/acp.1554>
7
8
9
10 Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie
11 nautical knots. *Learning and Instruction*, 14, 293–305.
12
13 <https://doi.org/10.1016/j.learninstruc.2004.06.005>
14
15 Seufert, T. (2003). Supporting coherence formation in learning from multiple representations.
16 *Learning and Instruction*, 13, 227–237. <https://doi.org/10.1016/j.learninstruc.2004.06.005>
17
18 Sheu, F. R., & Chen, N. S. (2014). Taking a signal: A review of gesture-based computing
19 research in education. *Computers & Education*, 78, 268-277.
20
21 <https://doi.org/10.1016/j.compedu.2014.06.008>
22
23 Stieff, M., Lira, M. E., & Scopelitis, S. A. (2016). Gesture supports spatial thinking in
24 STEM. *Cognition and Instruction*, 34, 80-99
25
26 <https://doi.org/10.1080/07370008.2016.1145122>
27
28
29 Tytler, R., Prain, V., Hubber, P., & Waldrip, B. (Eds.). (2013). *Constructing representations to*
30 *learn in science*. Springer Science & Business Media.
31
32
33 Videler, J. J. (1993). *Fish swimming*. London, UK: Chapman & Hall. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-94-011-1580-3)
34 [94-011-1580-3](https://doi.org/10.1007/978-94-011-1580-3).
35
36
37 Wiley, J. (2019). Picture this! Effects of photographs, diagrams, animations, and sketching on
38 learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology*,
39 33, 9-19. <https://doi.org/10.1002/acp.3495>
40
41
42 Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of
43 habit- formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Footnotes

¹ The German original version of the instruction read: “Während Du Dir die Videos ansiehst, möchten wir Dich bitten, mit Deinen Händen und/oder dem Körper Gesten auszuführen. Bitte gestalte die Gesten so, dass Du das Gefühl hast, dass sie Dir dabei helfen, das Lernmaterial besser zu verstehen. Es mag vielleicht komisch klingen, aber frühere Studien haben gezeigt, dass solche Gesten das Lernen verbessern und zwar vor allem dann, wenn sie gut auf das Lernmaterial abgestimmt sind.“ The term ‘gesture’ rather than ‘enactment’ was used since we found participants to better understand this term.

² We checked for homogeneity of regression slopes across conditions. There were no significant interactions between spatial ability and experimental conditions for any of the dependent variables. Thus, prerequisites for treating spatial ability as a covariate were met.

³ Importantly, even though medium and difficult tasks do not differ in terms of accuracy, the accuracy in more difficult tasks is achieved at higher temporal costs than that of medium-difficult tasks. Analyzing response times for the three task difficulty levels shows a linear increase in the mean response time per item from easy tasks ($M = 5705.56$ ms, $SE = 232.78$) to medium-difficult tasks ($M = 7894.30$ ms, $SE = 311.20$) to difficult tasks ($M = 9129.79$ ms, $SE = 336.03$) with significant differences between each level (all $ps < .001$).

Table 1

Means and standard deviations for students' entry characteristics as a function of experimental condition

	without enactment prompt		with enactment prompt	
	photographs	video	photographs	video
	(<i>n</i> = 22)	(<i>n</i> = 21)	(<i>n</i> = 20)	(<i>n</i> = 20)
Spatial ability (-10 to +10 points)	5.86 (3.00)	4.71 (4.95)	5.55 (3.24)	6.20 (3.19)
Verbal ability (max. 37 points)	29.14 (5.51)	28.76 (3.87)	29.05 (4.08)	28.55 (3.71)
Domain familiarity (max. 25 points)	6.61 (4.64)	6.88 (5.71)	7.13 (5.56)	6.65 (5.00)

Table 2

Means and standard deviations for mental effort, number of fishes with enactment and congruency as a function of experimental condition

	without enactment prompt		with enactment prompt	
	photographs (<i>n</i> = 22)	video (<i>n</i> = 21)	photographs (<i>n</i> = 20)	video (<i>n</i> = 20)
Mental effort (1-7)	4.98 (0.95)	4.08 (1.73)	5.00 (1.49)	4.83 (1.17)
Number of fishes with enactment (max. 6)	2.00 (2.73)	2.33 (2.76)	5.90 (0.45)	5.90 (0.31)
Mean Enactment Congruency (min. 0, max. 2)	1.31 (0.39)*	1.17 (0.32)*	1.19 (0.30)	1.46 (0.27)

*Note: based on *n* = 10 students performing enactments in these conditions

Table 3

Means and standard deviations for learning outcomes (0 - 1) as a function of experimental condition

	without enactment prompt		with enactment prompt	
	photographs	videos	photographs	videos
	(<i>n</i> = 22)	(<i>n</i> = 21)	(<i>n</i> = 20)	(<i>n</i> = 20)
Classification task				
Easy items	.92 (.12)	0.93 (.11)	0.94 (.10)	0.93 (.08)
Medium items	.64 (.23)	.61 (.21)	.76 (.18)	.71 (.22)
Difficult items	.64 (.19)	.67 (.19)	.68 (.19)	.74 (.19)
Description task	.69 (.13)	.59 (.19)	.72 (.18)	.62 (.18)



Figure 1: Sequence of stills for the *tetraodontiform* movement pattern (top row: 1-3, middle row: 4-6, bottom row: 5-9)

Acknowledgement: We thank Jana Arndt for her help in conducting the study.

*Highlights (for review)

- Can enacting non-human movements in your own body help learning?
- Participants studied videos or photos of fish locomotion; half were prompted to enact.
- Participants who enacted performed better on some classification tasks
- This was not influenced by the format of the visual materials
- Analysis of process measures suggests these results are due to general not specific mechanisms.

Credit Author Statement

Katharina Scheiter: conceptualization; methodology; analysis; writing

Birgit Brucker: methodology; analysis; investigation; project administration

Shaaron Ainsworth: conceptualization; methodology; analysis; writing