Does (Non-)Meaningful Sensori-Motor Engagement Promote Learning With Animated Physical Systems?

Wim T. J. L. Pouw[1], Charly Eielts[1], Tamara van Gog[1,3], Rolf A. Zwaan[1], and Fred Paas[1,2]

ABSTRACT— Previous research indicates that sensori-motor experience with physical systems can have a positive effect on learning. However, it is not clear whether this effect is caused by mere bodily engagement or the intrinsically meaningful information that such interaction affords in performing the learning task. We investigated ($N = 74$), through the use of a Wii Balance Board, whether different forms of physical engagement that was either meaningfully, non-meaningfully, or minimally related to the learning content would be beneficial (or detrimental) to learning about the workings of seesaws from instructional animations. The results were inconclusive, indicating that motoric competency on lever problem solving did not significantly differ between conditions, nor were response speed and transfer performance affected. These findings suggest that adult's implicit and explicit knowledge about physical systems is stable and not easily affected by (contradictory) sensori-motor experiences. Implications for embodied learning are discussed.

How does practical experience with physical systems (e.g., gear systems and levers) affect learning about its mechanisms? For example, does our *experience* with riding a bicycle contribute to our understanding of the working of mechanical parts (e.g., gear systems) of a bicycle? The answer to this question is relevant for educational practices in science education, as it dictates whether learning about physical systems should be grounded in concrete physical experiences next to abstract formalisms (Nathan, 2012; Pouw, Van Gog, & Paas, 2014). According to embodied learning theories, understanding of abstract principles relies upon the structural relations that emerge in bodily interaction with the environment (e.g., Goldstone & Barsalou, 1998; Lakoff & Núñez, 2000; Pouw et al., 2014). If this is correct, effective design of digital learning environments at times involves providing possibilities for bodily interaction.

Previous research indeed indicates that bodily interaction while learning or working with physical systems may in some cases promote understanding (e.g., Han & Black, 2011; Schönborn, Bivall, & Tibell, 2011; Zacharia, Loizou, & Papaevripidou, 2012). Such findings may prove informative for guiding applications of computer-based technology in education, such as tangible user interfaces (TUIs; e.g., Manches & O'Malley, 2012; Marshall, Price, & Rogers, 2003). TUIs are characterized by the combination of physical and virtual objects, running in real time, and allowing for physical interactions between the users and virtual objects that are typically afforded by interactions with real nonvirtual objects (Daponte, De Vito, Picariello, & Riccio, 2014). For example, the Nintendo Wii Balance Board can be used for continuous full-body *physical* interaction with *virtual* objects that can simulate complex affordances with nonvirtual objects (e.g., snowboarding). However, it is as yet unclear whether positive effects of bodily interactions with physical systems on understanding are promoted by the particular structural relations between agent and environment that emerge during physical interaction, or by the motivational processes that are affected by physical engagement (e.g., Bivall, Ainsworth, & Tibell, 2011; Han & Black, 2011; Wiebe, Minogue, Jones, Cowley, & Krebs, 2009).

---

[1] Department of Psychology, Education and Child Studies, Erasmus University Rotterdam
[2] Early Start Research Institute, University of Wollongong
[3] Department of Education, Utrecht University

Address correspondence to Wim T. J. L. Pouw, Department of Social Sciences, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands; e-mail: pouw@fsw.eur.nl.

In this study, participants learned about a class 1 lever (a lever where the fulcrum is in the middle and the effort and resistance on opposite sides; e.g., a seesaw) through physical engagement with a Nintendo Wii-Balance Board (hereafter: Wii Board).[1] Their physical engagement with the Wii Board was either minimally, meaningfully, or non-meaningfully related to the underlying principles of the physical system. This allows for studying not only *whether*, but also *how* bodily interaction supports learning of mechanical concepts.

Physical Engagement and Learning

There is increasing empirical evidence that physical engagement with learning materials can be an effective learning practice in for example mathematics, reading comprehension, and science education (e.g., Fyfe, McNeil, Son, & Goldstone, 2014; Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004; Kiefer & Trump, 2012). This research is informed by theories of embodied learning which suggest that learning and applying knowledge involves the effective reuse, simulation, or reactivation of sensori-motor experiences (for an overview, see Pouw et al., 2014).

Strong evidence for embodied learning comes from the field of gesture research, which has shown that actively producing, imitating, or enacting gestures during word learning (and retrieval) enhances memory (−retrieval) as opposed to more passive control conditions (e.g., Kelly, McDevitt, & Esch, 2009; Macedonia & Klimesch, 2014; for an overview see Macedonia & von Kriegstein, 2012). These findings are explained by the idea that the use of gestures during word learning enriches the conceptual understanding with multimodal information. This enrichment of the conceptual understanding is held to consist of a higher degree of associations with the concept's relevant modality-specific information (motor, haptic, spatial, etc.), which aids in the prevention of memory decay and the retrievability of the concept (e.g., Macedonia & von Kriegstein, 2012). Going beyond word learning, it has been found that gesturing (vs. not gesturing) during learning of science-related texts improves learners' ability to make inferences about the learning content (Cutica & Bucciarelli, 2013; Cutica, Ianì, & Bucciarelli, 2014).

A deeper analysis of the gesture literature shows that gesturing during learning is not effective merely by virtue of activating the sensori-motor system, but the meaningfulness of gestures appears to be important too. For example, it has been found that gesturing during word learning is only beneficial to memory when these gestures bear an iconic relation with the meaning of the word (e.g., moving hand up and down to depict "hammering") as opposed to gestures consisting of movements that are not concretely related to the semantic content of the word (e.g., Kelly et al., 2009; Macedonia & Knösche, 2011; see also Cook, Yip, & Goldin-Meadow, 2012). Thus, the gesture literature suggests that bodily activity might only aid learning when it is meaningfully related to the learning content.

Embodied Learning and Science Education

We are interested in how bodily activity might aid learning of principles underlying physical systems, a central learning topic in science education. To date, there are only a few quantitative experimental studies on the precise role of bodily engagement in this context (e.g., Bivall et al., 2011; Han & Black, 2011; Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014; Kontra, Lyons, Fischer, & Beilock 2015; Olympiou & Zacharia, 2012; Schönborn et al., 2011; Triona & Klahr, 2003; Wiebe et al., 2009; Zacharia & Constantinou, 2008; Zacharia & Olympiou, 2011; Zacharia et al., 2012). Next, we will only address findings that focus on concepts such as force and mass as these are central concepts for understanding the dynamics of a class 1 lever.

A demonstration of the benefit of recruiting sensori-motor processes in science concept learning is offered by a study of Bivall et al. (2011). They show that conceptual understanding of the structure of a bio-molecular model improved when learners were offered haptic feedback during the training phase. More precisely, learners were either engaging with a haptic device that simulated the repulsive and attractive forces of the molecules, or engaging with the same haptic device but with the haptic force-feedback disabled. Engaging with the haptic device with force-feedback bolstered the learning outcomes from pretest to posttest (for similar results, see Schönborn et al., 2011). It was suggested that haptic feedback during instruction afforded learners the opportunity to off-load visual working memory onto the sensori-motor system (Bivall et al., 2011; Schönborn et al., 2011). Furthermore, haptic feedback provided the learner with information about repellant and attractive forces *directly*, while that had to be visually inferred in the no-haptic feedback condition.

In a comparable study, fifth graders learned the workings of simple mechanical gear devices through different degrees of sensori-motor engagement (Han & Black, 2011). Subjects in the control condition observed the unfolding of the simulation, whereas in the other two conditions participants controlled the spinning of the gears with a joystick (kinesthetic condition); in the third condition, the joystick control was augmented with force feedback (force kinesthetic). Participants in the two kinesthetic conditions showed higher learning gains than participants in the control condition. Han and Black (2011) suggested that the kinesthetic experience allowed participants to reenact the relevant haptic information related to force as to actively compare it to visual information presented during the task.

In another study concerning workings of levers, it was investigated whether providing learners (ages 11–14) with haptic feedback during training would benefit learning performances as opposed to learners who only received visual information during training (Wiebe et al., 2009). In the first part of the training, participants set up the position of the lever's fulcrum and applied a number of weights. Subsequently, the program generated a second lever (with different position of the fulcrum and different number of weights). Participants were to judge which of the two levers (self-constructed vs. program-generated) required the highest amount of force to lift the weights. The participants in the haptic condition were also allowed to "feel" the amount of weight needed to lift the weights by using a device which produced haptic feedback. It was found that learning performance in terms of declarative or conceptual knowledge did not differ between the haptic and visual condition. In fact, participants in the visual condition outperformed those in the haptic condition in judging which lever required the highest amount of force.

There are indications, however, that the learning impact of physical engagement with objects or interfaces might be dependent on prior knowledge. For instance, in a study by Zacharia et al. (2012) kindergartners learned about the role of mass and its effects on a balance beam (class 1 lever) by either physically interacting with a balance beam or a virtual equivalent programmed on a computer. Prior to the training it was assessed whether children already possessed the correct conception that heavier objects placed on one side will pivot the balance beam. It was found that only children with an incorrect preconception benefited in terms of learning outcomes from physically interacting with a balance beam. This finding suggests that if learners already have an understanding of how mass relates to the balance beam they can assess mass based on perception alone and no additional sensori-motor information is needed to allow them to perceive mass directly (through kinesthetic feedback).

The previous results suggest that sensori-motor activity can be beneficial to learning underlying principles of physical systems. However, it should be noted that some of these studies did not find beneficial effects (Wiebe et al., 2009), and that some of the studies were very low powered (Bivall et al., 2011; Schönborn et al., 2011; Wiebe et al., 2009) to moderately powered (Zacharia et al., 2012); the study by Han and Black (2011) was an exception, it included a high number of participants. Even if we sidestep the issue of robustness of some of the previous findings, the design of the previous studies cannot rule out that bodily engagement only affects the motivational processes of the learner (e.g., Jones, Minogue, Tretter, Negishi, & Taylor, 2006). Most studies leave open the possibility that sensori-motor activity affects learning performance indirectly through affecting motivation and experiences of immersion, instead of by providing *meaningful* information about the learning content. As research on gesture and learning shows, it appears likely that only meaningful physical engagement would promote learning, but it cannot be ruled out that the physical engagement as such (and the structural relations that are picked up) had indeed benefited learning in those studies. Therefore, the aim of the present study was to address *how* enriching the learning content with sensori-motor information affects learning.

The Present Study

To study *how* enriching the learning content with sensori-motor information affects learning, we manipulated the meaningfulness of the structural relations between physical actions on the Wii Board and the instructional animations of the learning content (i.e., mechanical principles of class 1 levers; a seesaw). Participants were assigned to one of three conditions. In the first, subjects were given a meaningful embodied training, in which they learned to balance a seesaw across several trials, by applying force on the Wii Board that matched the number and position of the weights that acted as counterforce on the seesaw (*meaningful condition*). In the second condition, subjects were given a similar training, but in this condition the forces that needed to be applied on the Wii Board to balance the seesaw were non-meaningfully correlated with the number and position of the weights that acted as counterforce (*non-meaningful condition*). Yet, in this condition participants did apply force on the congruent side of the seesaw. Thus, while participants pushed a seesaw down on the congruent side, the force needed to push the seesaw into balance was not consistently (i.e., non-meaningfully) related with the number of weights placed on the seesaw. These conditions were compared with a third, *minimal condition*, in which participants merely provided a small push that started an animation of a seesaw balancing out. Thus, importantly, participants in all three conditions are using the Wii-Board to interact with the instructional animation, which allows us to eliminate some of the motivational effects on performance that might arise from the mere use of the Wii-Board and from having the animation respond to an action by the learner.

Not only do we explore whether meaningful physical experiences may support learning, we also assess whether non-meaningful physical experiences (i.e., acting with incorrect relations with the learning principle) hamper learning. After all, if knowledge is indeed grounded in action as embodied theories of learning have it, then we might also predict the opposite, namely physical experiences that are incongruent with the learning principle should hamper learning. This is a novel question that allows us to further gauge the degree to which knowledge of physical systems (i.e., levers) is affected by sensori-motor experience.

To assess the broad aspects of learning afforded by sensori-motor interaction, we used three different performance measures: a reaction time task (henceforth RT task), a transfer task, and a motor task. The RT task relied heavily on visual perceptual experiences; it assessed speed and accuracy of judging whether a depicted seesaw should balance out or pivot given the weights and their position. The transfer task relied more on deliberate reasoning; it measured the accuracy of judgments about more complex class 1 lever concepts (e.g., interconnecting seesaws and varied fulcrum positions). The motor task relied purely on motoric knowledge; it assessed whether participants were able to physically enact the correct amount of force to balance a seesaw when provided with a noninteractive picture of a seesaw.

This motor task provides us with a novel and exploratory way to assess whether knowledge of mechanical systems can be partly assessed in the way subjects enact the solution of the problem as opposed to tasks that are procedurally very different in nature (pushing a button; i.e., RT task and the transfer task). Essentially, it allows us to assess whether our learning manipulations differentially affect whether participants *know how to* physically balance a seesaw (motor task) as opposed to *knowing that* a seesaw balances out under particular conditions (RT and transfer task; Ryle, 1945).

To assess cognitive load and motivation differences, we also included subjective attitudes (mental effort, interest, difficulty) toward the learning phase and test phases to check for possible mediating effects of motivation (interest) and experienced difficulty. Participants' reports of the interest of the learning phase are of special concern to the present study, as they provide a way to assess whether there were motivational differences across conditions.

We hypothesized that participants in the meaningful condition would outperform participants in the minimal and non-meaningful conditions on all performance tests. We also hypothesized that the non-meaningfully embodied instructional animation (i.e., non-meaningful condition) would actually hinder performance on these tasks as compared to the other conditions, as it provides interfering sensori-motor information.

METHOD

Participants and Design

A total of 92 Dutch university students participated in the present study for course credit or 10 euros. Unfortunately, due to a programming error for 15 participants the Wii-Board data were lost (meaningful = 4, non-meaningful = 5, minimal = 6). Additionally, one participant (non-meaningful) was excluded from the analyses for not following the instructions correctly (participant employed two hands instead of one to push on one side of the WiiBoard). This resulted in data of 76 participants for the analyses (37 males [48.68%]; age range = 18 to 25, $M = 21.32$, $SD = 2.112$; 93.4% right handed, as determined by Oldfield, 1971), who were randomly distributed among three conditions in a between-subjects design: meaningful ($N = 26$), non-meaningful ($N = 25$) or minimal ($N = 25$).

Materials

Instructional Animations

The voice-over and textual instructions and self-report questions were programmed in ActionScript 3.0 and the animations were designed in Adobe Flash Professional CS 5.5 (see http://www.charlyeielts.nl/wbb/materials.html or https://osf.io/ebjvm/). The Wii Board communication was handled by the WiiFlash Actionscript API and WiiFlash Server developed by Joa Ebert and Thibault Imbert (http://wiiflash.bytearray.org/).

Prior to this study we assessed whether adults were affected in performance in one of our main learning measures (reaction time task) by comparing the effect of only observing the instructional animation as opposed to receiving no instructional animation. This was to ensure that adults are still receptive to training about class 1 levers. In this pilot study with adults ($N = 78$; 52.6% female; age $M = 33.47$, $SD = 12.29$, with 83.4% reporting having had college experience) using Amazon's Mechanical Turk we used the exact instructional materials designed for this study but without possibilities for physical interaction. This pilot study showed that the animations were effective for learning (57.24%, $SD = 19.4\%$ accuracy on the reaction time task) as compared to no instruction (69.26%, $SD = 20.4\%$), $t(76) = -2.644$, $p = .010$, Cohen's $d = .602$ [large effect]). No effects were obtained for solving speed on the RT task, $t(76) = -0.945$, $p = .348$, Cohen's $d = 0.218$.

Introductory instructional animation. Before the manipulation phase, each participant viewed a short noninteractive instructional animation of 190 s with a Dutch female voice-over, in which the different concepts involved in the operation of a lever were introduced (introduction phase). The introduction phase presented the seesaw and its components (fulcrum, left arm, and right arm), and the concepts of load, force, and balance. This introduction phase further focused on the mechanical principle of levers. The mechanical advantage principle explained in this animation involved the concept that force can be amplified by increasing the distance from the fulcrum.

Manipulation: Interactive instructional animation. In the manipulation phase, participants had to perform 24 interactive study trials in which they had to return a tilted seesaw to a state of balance using the Wii Board.[2]

Before each of these trials, a fixation cross was displayed and subjects were instructed not to apply any force on the Wii Board. The experiment would automatically start when the subjects employed force that did not deviate more than 0.2 lbs from the calibration values for longer than 500 ms, which ensured that every study trial started from a rest position. At the beginning of each trial, a seesaw was presented that could be divided into nine even-sized parts with the fulcrum placed in the middle. In each trial, the seesaw was either tilted left or right with one weight (either small [one cube] or large [two cubes]) placed on one side of the seesaw. In half of the trials a load of one blue cube was tilting the seesaw and in the other half, a larger load of two blue cubes stacked on top of each other tilted the seesaw. The animation was designed such that the large weight was exactly two times the volume of the small weight. Participants were instructed to return the seesaw to balance by employing a required amount of force on the opposite arm of the seesaw. The location on the seesaw where the required force should be applied was marked by a yellow highlight around the edges of the area. When subjects applied the correct amount of pressure on the Wii Board, the seesaw would react and the arm carrying the counter weight would be lifted from the ground and the cube representing the participant's administered force would grow to the correct size to establish balance. The animation would stop if a state of balance was reached.

The required force to balance a seesaw differed across conditions. For the meaningful condition, the required pressure for small weights was 5 lbs, with a range of 4 to 6 lbs and 10 lbs for large weights with a range of 9 to 11 lbs. In the non-meaningful condition, the force requirements of 5 and 10 lbs were *randomized* for the small and large weights, so that there was no structural correlation between amount of weight and counterweight to achieve balance across trials. In both the meaningful and non-meaningful condition, the seesaw would go out of balance at the force side if the upper bound of the accepted range of employed force was exceeded. If the applied force was lower than the required minimum, the seesaw would return to its initial state. In the minimal condition, the animation would simply play if participants prompted it to start by shortly applying a small amount of pressure (>0.3 lbs) on both sides of the Wii Board. Importantly, when the seesaw was in balance, participants in the meaningful and non-meaningful motor conditions had to continue employing the appropriate amount of force for 2 s before the experiment proceeded to the next trial. In the minimal condition, the experiment would automatically proceed to the next trial 2 s after the seesaw reached a state of balance.

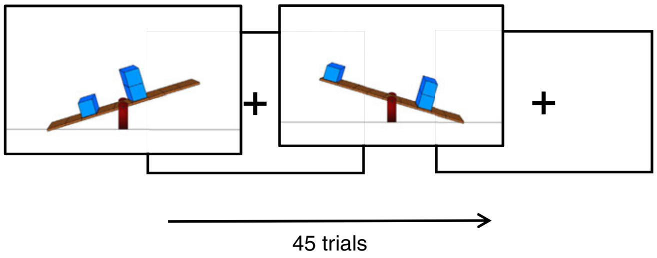

Fig. 1. Example of two reaction time trials. After each response a fixation cross would appear intermittently.

Test Tasks

Reaction time task. We developed a three-choice reaction time task programmed in E-Prime (henceforth RT task) to assess participants' accuracy (number of correct responses) and efficiency (reaction time) in assessing class 1 lever's mechanics. In this RT task, participants were shown a seesaw that was either in balance or tilted to the left or right. In each trial, one or two blocks are presented on each side of the seesaw on deferring distances from the fulcrum. The size and location of the weights varied across the 45 trials. Subjects had to determine which way the seesaw should be tilted given the presented weights, regardless of the current state of the seesaw (i.e., pivoted to left/right or balanced). Subjects responded with a keyboard by pressing *P* if the seesaw should be tilted to the right, *Q* if it should be tilted to the left, and *SPACE* if the seesaw should be in balance. Subjects were instructed to respond as fast as possible. Thirty-two trials of the 45 consisted of a situation where the principle of mechanical advantage was relevant, meaning a weight was closer or further from the fulcrum than the opposite weight (see Figure 1 for an example).

Transfer task. The transfer task consisted of a total of 12 trials consisting of a three-choice judgment task. Participants were prompted to think as long as they needed to produce the correct answer. The trials required participants to judge whether a seesaw in a set of several interconnected see saws and differing positions of the seesaws' fulcrum, would pivot to the right, to the left, or would stay balanced (see Figure 2). Also four trials involved the judgment of the amount of force needed to balance two seesaws in which participants had to judge which arm of the avatar needed to exert the most amount of force to balance the seesaws.

Motor task. In the *motor* task, participants had to determine the amount of force that needed to be employed to balance a seesaw that was statically presented, by applying the force on the Wii Board (22 trials). These trials were identical to the practice trials in the study phase, with the exception that the seesaw could not be controlled via the Wii Board (i.e., it remained a static picture). During each

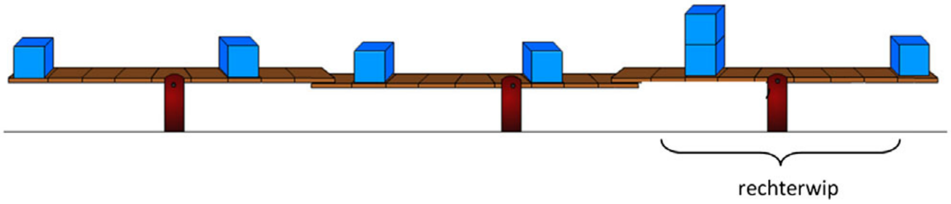

Fig. 2. Example of one transfer task trial. Participants were asked to "Judge whether the right seesaw (*rechterwip*) pivots to the left, remains in balance, or pivots to the right."

trial, participants would employ force for 1500 ms. During the trials we assessed the amount of weight applied by the participant over a 1,000 ms period (sampling: 60Hz) for 22 trials and sampled force values from 500 ms onward (thus 500–1500 ms). We decided not to sample the first 500 ms of force employment as we were interested in the moment participants reached a stable force employment. This allowed us to gauge participants' ability to correctly judge the different levels of force that should be employed for balancing the seesaw.

Self-Report Questions

Mental Effort, Difficulty, Interest. As an indication of experienced mental effort, perceived difficulty, and experienced interest during the learning phase and after each of the test phases (RT task, transfer task, and motor memory task) participants answered on a 5-point scale "How much mental effort did you invest during the learning phase (or RT task, transfer task, motor task)" (mental effort; 1 = "very low mental effort", to 5 = "very high mental effort"), "How difficult did you find this task" (difficulty; 1 = "not difficult", to 5 = "highly difficult"), and "How interesting did you find this task" (interest; 1 = "not interesting", to 5 = "very interesting").

Physical Effort. Amount of physical effort invested in completing the interactive animation trials ("how much physical effort did you exert during this task" on a 5-point scale (1 = "very low effort" to 5 = "very high effort").

Handedness. Using a modification of the Oldfield (1971) Handedness questionnaire participants reported hand dominance for several manipulative situations (e.g., writing, brushing teeth, etc.) on a 5-point scale (1 = very left hand dominant, to 5 = very right-hand dominant). We computed the mean responses and categorized left (right) handedness for means lower (higher) than 3.

Prior Knowledge Self-Report. Prior knowledge of the learning material ("Before this experiment, I was knowledgeable about levers"; 1 = not knowledgeable, to 5 = very knowledgeable). We also checked whether participants had a physics background obtained in secondary school (0 = no, 1 = yes).

Demographics. Participants reported after the experiment their age, gender, and study program, and were allowed to comment on the nature of the experiment.

Procedure

Participants were informed that they would start with a training phase and would subsequently perform several learning tests. First, participants were seated at a table on which a Wii Board was mounted. The chair's height and elbow supports were adjusted such that participants' hands rested on the Wii-board and the elbows had a 90-degree angle. The Wii Board was first calibrated to the participant's resting state. During the calibration procedure, participants positioned their left and right hand on the corresponding side of the Wii Board with their hand placed on a marker that represented the location of the pressure sensors in the Wii Board for either side. It was stressed that the subjects should only rest their hands on the Wii Board and should not apply any pressure on the Wii Board during the calibration. In order to familiarize participants with the Wii Board controls, they performed a short training sequence before the experiment. The training consisted of four different cubes that changed color if the participant gave the correct amount of pressure on the Wii Board. The different levels of pressure corresponded to the ones used in the experimental procedure. If the pressure exceeded the required force, the color of the cube would overflow and participants were instructed to apply less pressure. Participants then watched the introductory instructional animation and subsequently proceeded to the interactive instructional animation, which they interacted with in a meaningful, nonmeaningful, or minimal manner depending on their assigned condition. During the training phase, the study time (time needed to balance the seesaw in the 24 trials) was recorded by the software, as this was likely to vary across conditions as a result of the experimental manipulation. After this interactive training phase, participants reported exerted physical effort. Subsequently, participants performed the RT task, transfer task, and motor task (in that order). After the training phase and test phases, participants reported their experienced mental effort, difficulty of task, and interest in task. Finally, participants answered questions concerning gender, age, prior knowledge, handedness, and physics background as reported above.

Data Analyses

Accuracy and RT scores for the transfer task and RT task lying outside 2.5 *SD* of the overall mean were replaced with the overall mean (will be reported in the results if applicable).

Reaction time task. The number of correct answers on 45 trials (performance range: 0–45) was taken as a measure of accuracy and the mean reaction time (in ms) on correct trials as a measure of speed.

Transfer task. The number of correct answers on 12 trials was taken (performance range: 0–12).

Motor task. We obtained two outcome measures from the motor task. Firstly, we provide the different trajectories for the applied force during 1,000 ms for the two different levels of force (one cube vs. two cubes to balance a seesaw). This should give us exploratory information about whether the conditions indeed performed differently, as can be expected since participants learned to balance a seesaw with differing weights. Because we are interested in whether participants' motor performance reflects understanding of the mechanics of a seesaw, we used an additional *ratio measure* which reflects whether participants could correctly differentiate between one versus two cubes, that is, one cube should be half the force of two cubes. This was done by dividing the mean amount of force given for one-cube trials (11 trials) by the mean amount of force for two-cube trials (11 trials); when participants indeed were able to correctly differentiate between one versus two cubes the ratio would give $1/2$ value (i.e., .5). The final measure is therefore the absolute difference of the correct ratio of .5 and the ratio attained by the participants; $\left[\frac{\text{Mean Force Cube 1 Trials}}{\text{Mean Force Cube 2 Trials}} - .5\right]$; this yields .0 as a perfect score (i.e., lower score is better).

Unfortunately, due to technical issues we failed to administer Wii-board data for this particular task for an additional seven participants, yielding a sample of 68 participants (meaningful [$N = 23$] vs. non-meaningful [$N = 24$] vs. minimal condition [$N = 21$]).

RESULTS

Prior Knowledge and Physics Background

Using analysis of variance (ANOVA), no significant differences (see Table 1 for means) were found across conditions for prior knowledge, $F(2, 73) = 2.27$, $p = .110$. This was also the case for physics background, $F(2, 73) = 1.078$, $p = .346$.

Wii-Board Training Phase

Training Duration

The duration of the training phase differed, such that the non-meaningful condition ($M = 184.50$ s, $SD = 107.15$) was longer in duration than the meaningful condition ($M = 133.79$ s, $SD = 27.56$) and the minimal condition ($M = 95.12$ s, $SD = 73.26$). As is evident, variances were not equal across groups (Levene's $\alpha < .001$). To test whether the differences in training-phase duration were significant, we performed a Kruskal–Wallis analysis with pairwise comparisons. There was a significant overall effect of condition on duration, $\chi^2(2) = 37.519$, $p < .001$. Pairwise comparison showed that the non-meaningful condition and the meaningful condition took longer than the minimal condition (minimal vs. non-meaningful condition, χ^2 $[1] = 5.991$, $p < .001$; minimal vs. meaningful condition, $\chi^2[2] = 4.126$, $p < .00$). However, the meaningful condition did not differ from the non-meaningful condition, $\chi^2(1) = -1.924$, $p = .163$.

Task Load

See Table 1, column 1, for the means and standard deviations for the reported mental effort, difficulty, and interest for the training phase across conditions.[3] One-way ANOVAs only showed a significant effect of condition on difficulty, $F(2, 73) = 11.754$, $p < .001$, $\eta_p^2 = .24$. Post hoc comparisons (Bonferroni) showed that the non-meaningful training phase ($M = 2.68$) was reported to be significantly more difficult than the minimal training phase ($M_{\text{difference}} = -1.20$, $p < .001$), and the the meaningful training phase ($M_{\text{difference}} = -.72$, $p = .014$). The meaningful training phase did not differ on difficulty from the minimal training phase ($M_{\text{difference}} = -.48$, $p = .164$).

Physical Effort

Reported physical effort during the training phase in the minimal condition ($M = 1.88$, $SD = 0.93$) the meaningful ($M = 2.35$, $SD = 1.06$), and the non-meaningful ($M = 2.56$, $SD = 1.16$) did not differ, $F(2, 75) = 2.738$, $p = .071$).

RT Task

Accuracy

We replaced outliers (outside 2.5 *SD* range from the mean) with the overall mean ($n = 2$). Overall accuracy was 80.04% ($M = 36.02$ correct responses out of 45, $SD = 3.26$), with the meaningful condition scoring 80.05% ($M = 36.23$ [/45], $SD = 2.83$), the non-meaningful condition scoring 79.56% ($M = 35.80$ [/45], $SD = 3.30$), and the minimal condition scoring 80.00% ($M = 36.00$, $SD = 3.73$); also see Figure a. A one-way ANOVA yielded no significant differences across conditions, $F(2, 75) = .109$, $p = .897$, $\eta_p^2 = .003$. An additional

Table 1
Means and Standard Deviations for Mental Effort, Difficulty, and Interest for Training Phase and Performance Tasks Across Conditions

Condition		*Training Phase*		*Reaction Time Task*		*Transfer Task*		*Motor Task*		*Prior Knowledge*		*Physics Background*	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Meaningful condition	Mental effort	3.15	0.68	3.19	0.85	3.73	0.92	2.73	1.22	3.27	1.19	.231	0.07
	Difficulty	1.96	0.92	3.73	0.72	3.88	0.86	3.73	1.185				
	Interest	1.62	0.85	3.62	0.80	3.81	0.80	3.35	0.80				
Non-meaningful condition	Mental effort	2.80	0.96	3.32	0.95	3.80	0.76	2.72	1.17	2.60	1.12	.160	.07
	Difficulty	2.68	1.03	3.56	0.96	3.76	1.05	3.16	1.179				
	Interest	1.64	.64	3.28	0.84	3.81	0.81	2.96	0.98				
Minimal condition	Mental effort	2.68	0.99	3.04	0.94	3.28	1.06	2.20	0.96	2.80	1.16	.08	.07
	Difficulty	1.48	0.65	3.56	1.08	3.56	1.08	3.20	1.291				
	Interest	1.36	0.70	3.52	0.77	3.84	0.69	2.96	0.98				

Bayesian analysis for the effect of motor-involvement condition on accuracy yielded $p_{BIC}(H_0|D) = .986$ (Masson, 2011). This probability indicates a 98.6% likelihood that motor-involvement condition (meaningful, non-meaningful, and minimal) does not affect accuracy on the RT task. Following guidelines by Kass and Raftery (1995), this information criterion is strong evidence for the absence of an effect of motor involvement on accuracy.

Furthermore, in our pilot study with adults on Mechanical Turk, those participants ($N = 43$) who did not view an instructional animation had a considerably lower accuracy score on the RT task (57.24%) than participants ($N = 35$) who did view animations (but without opportunities for interaction) in the pilot study (69.26%) and participants in the present study where overall accuracy was 80.04%. In sum, the pilot study data suggest that the instructional animations used here contribute to learning.

Reaction Time

No outliers outside the 2.5 SD range from the mean were found. The average reaction time in *ms* for correct trials (see Figure 3b) for the meaningful condition ($M = 2152.83$, $SD = 489.45$), the minimal condition ($M = 2396.02$, $SD = 857.29$) and the non-meaningful condition ($M = 2717.07$, $SD = 1202.200$) showed unequal variances across condition (Levene's $\alpha < .001$). We performed a Kruskal–Wallis analysis with pairwise comparisons which yielded no significant overall effect of condition on reaction times, $\chi^2(2) = 1.860$, $p = .395$.

Task Load

See Table 1, column 2, for the means and standard deviations for the reported mental effort, difficulty, and interest on the reaction time task across conditions. ANOVAs showed no significant overall main effects of condition on these self-report measures regarding the RT task. Furthermore, there were no significant correlations of self-report measures with performance on the RT task.

Transfer Task

Accuracy

No outliers outside 2.5 SD range from the mean were found. Overall mean accuracy (see Figure 4) was 52% (i.e., a mean of 6.24 correct responses out of 12, $SD = 1.6$); meaningful ($M = 6.54$ [%], $SD = 1.363$), non-meaningful ($M = 5.84$ [50.00%], $SD = 1.625$), and minimal condition ($M = 6.32$ [50.00%], $SD = 1.77$). A one-way ANOVA yielded no significant differences across condition, $F(2, 73) = 3.24$, $p = .258$, partial $\eta_p^2 = .034$. An additional Bayesian analysis for the effect of motor-involvement condition on transfer task accuracy yielded $p_{BIC}(H_0|D) = .954$ (Masson, 2011). This probability indicates a 95.4% likelihood that motor-involvement condition (meaningful, non-meaningful, and minimal) does not affect accuracy on the transfer task; which can be considered strong evidence for the absence of an effect of condition (Kass & Raftery, 1995).

Task Load

See Table 1, column 3, for the means and standard deviations for the reported mental effort, difficulty, and interest on the transfer task across conditions. ANOVAs showed no significant overall main effects of condition on these self-report measures regarding the transfer task. With regard to overall correlations between self-report measures and performance on the transfer task, we only found a significant correlation between experienced interest and accuracy on the transfer task, such that more reported interest resulted in higher performance, $r = .292$, $p = .011$.

Motor Memory Task

Figure 5 shows the mean force responses for one-cube (gray) and two-cube trials (lock) plotted over time (1 s) with 95%

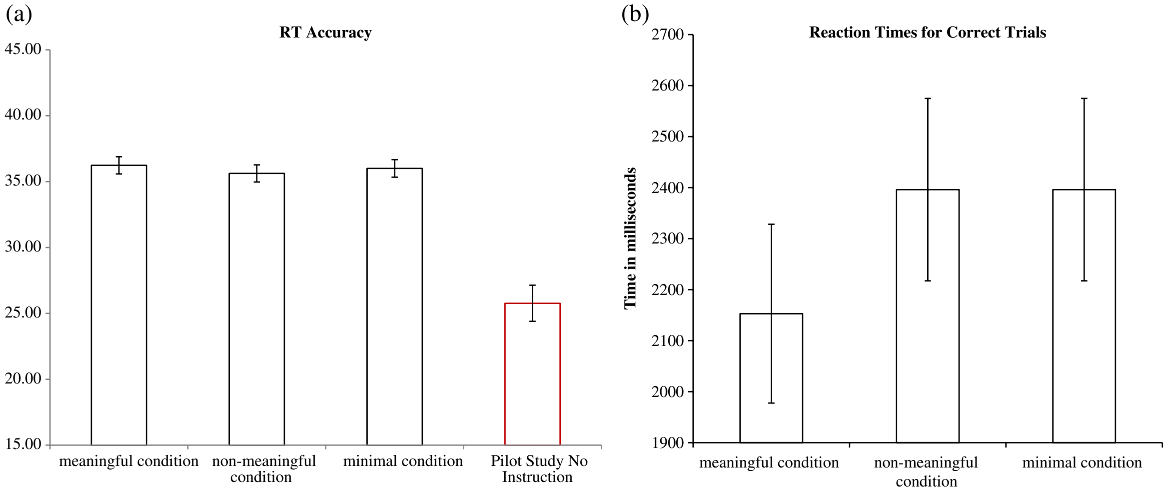

Fig. 3. (a, b) Number of correct trials and reaction times for correct trials (ms) for the meaningful, non-meaningful, and minimal conditions. Error bars indicate standard errors. For number of correct trials (RT accuracy, panel a) we have added the results of a pilot study performed on Mechanical Turk that shows the accuracy for the group that did not receive instructional animations.

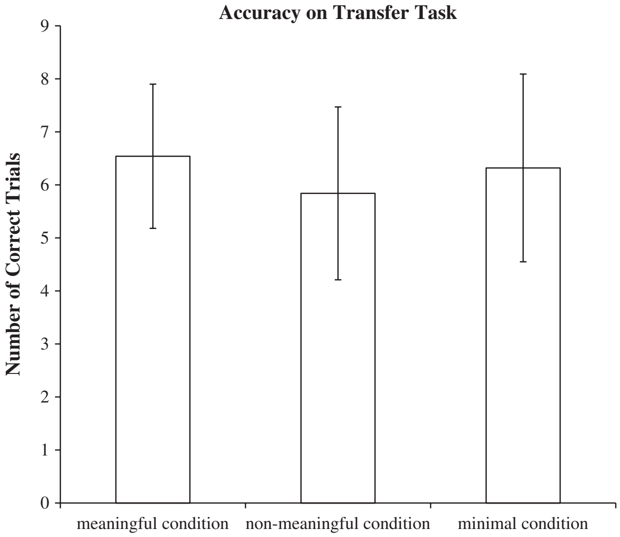

Fig. 4. Number of correct responses on the transfer task for the meaningful, non-meaningful, and minimal conditions. Error bars indicate standard errors.

confidence intervals. As can be qualitatively inferred from the graphs, there is a considerable difference in participants motor responses in the minimal condition as compared to other conditions. This is not surprising, as participants in the minimal condition were trained to give a short push on the Wii-Board to balance the seesaw, which is reflected in this task as well. Namely, participants in the minimal condition gave a short but large force response as compared to the other conditions.

However a more interesting pattern appears to have emerged if we consider that only participants in the meaningful condition were trained to motorically differentiate between forces of one versus two blocks because the forces in the non-meaningful condition were not consistently related to the weights. Further consider that participants in the control condition only gave one force response with both hands that did not covary with one versus two blocks. Interestingly, the figures appear to indicate that, indeed, the non-meaningful condition motorically differentiated less between one versus two blocks as compared to the meaningful condition. Moreover, the participants in the minimal condition—although not motorically trained to differentiate between weights—did appear to transfer their knowledge motorically, as indicated by the distances between curves.

Ratio Measure

To test whether these differentiations for force responses for one- versus two-cube trials were significant we obtained a ratio measure as described in the Method section. As is shown in Figure 6, the participants in the meaningfully condition ($M = .3026$, $SD = .156$) performed p (a score of 0 being perfect) in differentiating between one-cube- versus two-cube forces as compared to the non-meaningful ($M = .362$, $SD = .125$) and minimal conditions ($M = .354$, $SD = .163$). However, a one-way ANOVA yielded no significant differences across condition, $F(2, 78) = 1.091$, $p = .342$, partial $\eta_p^2 = .032$. An additional Bayesian analysis for the effect of motor-involvement condition on ratio measure on the motor memory task yielded $p_{BIC}(H_0|D) =$ .956 (Masson, 2011). This probability indicates a 95.6% likelihood that motor-involvement condition (meaningful, non-meaningful, and minimal) does not affect motor

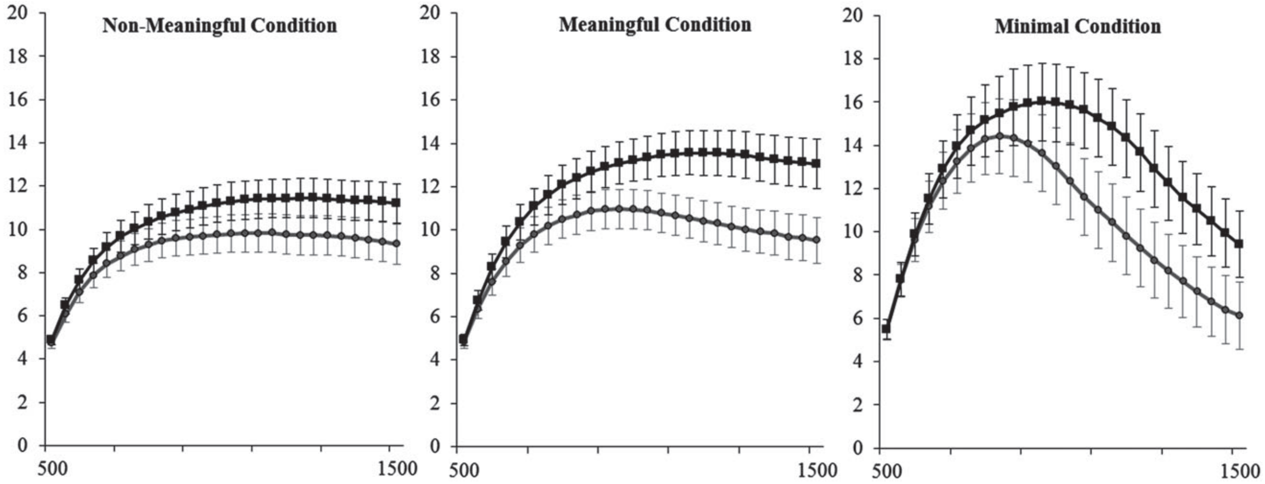

Fig. 5. The force responses for the motor memory task over time (1,000 ms) for trials (differentiated by force-for-one-cube [in gray] and force-for-two-blocks [in black]) per condition. Error bars indicate 95% confidence intervals.

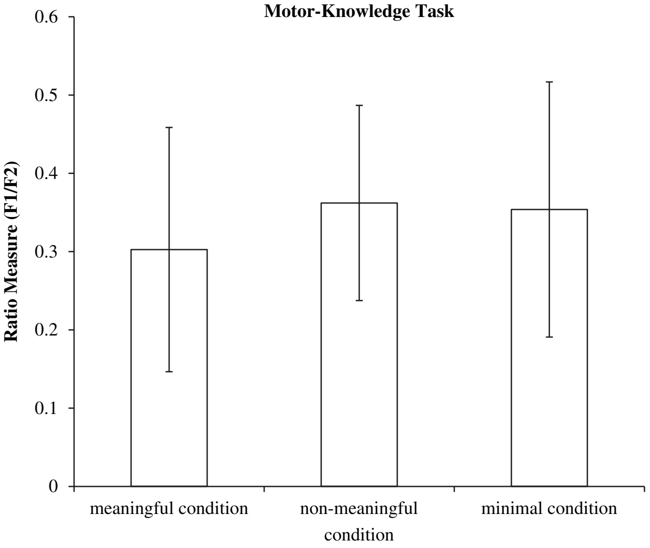

Fig. 6. Ratio measures as a measure of motor competency to differentiate between force-for-one-cube versus force-for-two-cubes trials. A score of 0 means a perfect score, meaning that participants' mean force given for two cube trials was twice the force compared to one cube trials.

knowledge as reflected by the ratio measure (which can be considered strong evidence for the absence of an effect of condition; Kass & Raftery, 1995).

Task Load

See Table 1, column 4, for the means and standard deviations for the reported mental effort, difficulty, and interest of the motor task across conditions. ANOVAs showed no significant overall main effects of condition on these self-report measures regarding the motor task. Overall correlations between self-report measures and performance showed that those who found the task more difficult ($r = -.267$, $p = .028$) and more interesting performed ($r = -.263$, $p = .030$) better on the motor task.

DISCUSSION

We investigated whether different meaningful- and non-meaningful forms of physical engagement with instructional animations concerning the workings of class 1 levers affects unreflective (RT task), reflective (transfer task), and motoric (motor task) competency on problem solving. The results showed that either a training phase in which participants learned how to physically balance a virtual seesaw (meaningful condition), a training in which participants physically balance a seesaw but with inconsistent weight mapping (non-meaningful condition), or a training phase in which participant merely activated the playing of an instructional animation through a minimal physical engagement (minimal condition), did not differently affect performance on RT task or the transfer task.

Participants in the minimal condition did have different motoric judgments of the force that needed to be applied to balance a seesaw. This result is not surprising, because participants in the minimal condition learned to balance a seesaw only through minimal physical engagement (3 lbs) whereas the other participants consistently or inconsistently learned to balance a seesaw around 5 and 10 lbs for one-cube and two-cube forces, respectively. Nevertheless, this result confirms that there was *some* implicit embodied memory of the correct sensori-motor dynamics with the seesaw during the training phase. Yet, no significant differences were found between conditions for the ratio measure, which was designed to assess motoric competence in correctly differentiating between one- versus two-cube forces over an interval of 500–1,500 ms. Interestingly, as the visual evaluation of Figure 6 shows, participants in the minimal condition show *some* motoric competence (as indicated by differentiation of forces between one versus two blocks between 1,000 and 1,500 ms after response onset), which suggests that knowledge about the mechanisms of the seesaw learned through nonmotoric means may transfer to motoric competence.

Yet, this study has some limitations that might have prevented us to find the hypothesized beneficial (negative) effect of the meaningful (non-meaningful) training phase as compared to the minimal condition. First, although the current paradigm was explicitly designed to pick up potential small effects of training of a nonreflective and automatic sort, it might be the case that the manipulation was simply too short to imbue effects of the different training phases (approximately 2 min). Indeed, it could be argued that perhaps especially in learning sensori-motor routines repetition is important to achieve a certain level of competence (e.g., Marley & Carbonneau, 2014). This can be appreciated by the idea that, in contrast to understanding a propositional rule, motor competence does not follow an *either-or* transition of understanding (cf. Ryle, 1945). Thus, for the learner, practice might be a very important factor to pick up information that is constituted by the structural correlations that emerge during interaction; or, in simpler words, embodied learning takes time.

Another limitation of the present design is that we could have obtained a more sensitive measurement by including a pretest. For example, Zacharia et al. (2012) showed that children that had correct conceptions of mass and its effect on a balance beam were not benefiting from physically engaging with learning materials. It is thus possible that the learners' degree of competence affect whether physical engagement is beneficial for learning; unfortunately the present design fails to take this into account.

Additionally, it might be argued that the null findings actually show that the more passive training (i.e., minimal condition) was more efficient for learning than the other forms of physical engagement. After all, participants in the minimal condition had a significantly shorter study time as compared to the meaningful and non-meaningful conditions. Unfortunately this is difficult to assess. However, the reason why the embodied instructional animations took longer is that participants had to acquire competence in wielding the Wii-board (e.g., during the training phase participants often over-pushed and then stopped pushing altogether to begin all over again). As such it can be argued that participants in the physically engaged conditions were actually performing several tasks at once, and were thus in another respect hindered to study the materials.

Methodological issues aside, given that previous research (with children and adults) does not consistently find a potential beneficial role of augmenting instructional animations with sensori-motor information (Bivall et al., 2011; Schönborn et al., 2011; Wiebe et al., 2009; Zacharia et al., 2012), it might be the case that learning *how* to do something physically is not always necessary to know *that* a mechanical device works such and so. In other words, perhaps learning the workings of levers can be done entirely through visual information alone in the current task (i.e., the actions participants performed in the meaningful condition were not relevant). In fact, it might have worked the other way around. *Knowing that* informs how to motorically balance a seesaw, as indicated by the apparent motor competence of participants in the minimal condition. Indeed it has been argued that when visual information is present and usable to understand a particular task at hand, haptic information—even when it provides extra information—will not necessarily be used next to visual information (Driver & Spence, 1998; Klatzky & Lederman, 2002). Moreover, it may be the case that integrating haptic information with visual information produces additional cognitive load which counteracts potential beneficial effects of extra-visual information provided by haptic interaction (Skulmowski, Pradel, Kühnert, Brunnett, & Rey, 2016). Yet, it is important to note that on our reading of most theories on embodied learning (for an overview of such theories see Pouw et al., 2014) the current actions performed with the virtual seesaw would be relevant for further reasoning with seesaws. Namely, learning to judge the force needed to balance a seesaw motorically corresponded lawfully (in the meaningful condition) with the visual information (i.e., the number of blocks, in combination with the position of the blocks on the arms of the seesaw, lawfully corresponded to the force that needed to be applied by the participant on the relevant arm). Although of course it cannot be excluded that a more natural correspondence of action and perception (say interacting with an actual seesaw) would have provided different results. Nevertheless, the visual information presented in the motoric training sessions directly corresponded with the visual information provided in the subsequent performance RT and transfer tasks. Embodied learning theories prescribe that after motoric experiences further visual encounters with similar situations are laden with previous multimodal associations and become in fact part of reasoning with such visual information (e.g., Barsalou, 1999). If the current data reflect a true null effect, it thus appears to suggest (for the present context) that these multimodal associations predicted by embodied learning theories are (1) either not established after a short motoric training or/and thus (2) not used for further reasoning.

Additionally, it might be that basic concepts such as weight and mass are learned early on in childhood, and thus need not be provided with extra information any more. In other words, certain basic concepts are already grounded in physical experiences. For example, in the current task, because participants were able to differentiate between one cube and two blocks separately, as well as the position of the blocks through visual information alone, it might have rendered the motoric information redundant for the participant. As such, grounding science content in physical experiences is only necessary if it adds something otherwise unknown to the learner (e.g., Pouw et al., 2014; Zacharia et al., 2012). However, the idea that participants' knowledge is still grounded in physical experiences is less informative to address the current results given the finding that participants were not affected by contradictory physical experiences provided in the non-meaningful training phase. Furthermore, we should highlight (as reported in the Method section and in our results) that we performed a pilot test using instructional animations about class 1 lever problems wherein we did find large performance effects of animation versus no animation on the RT task with adults. Additionally, as reported in the Results section, in the pilot study it was already established that providing participants with a similar but noninteractive instructional animation leads to better performance (69.26%) on the RT task than no animation (57.24%). If we consider that as a baseline for the current study (see Figure 3a) in which participants interacted with the animation, we see that the present sample performs much better on that same task (80.04%). This suggests that the instructional animations are effective for improving performance compared to no training, and therefore that the current lack of differences between conditions cannot be explained merely by poor learning effectiveness (i.e., floor effect) of the instructional animations. Furthermore, with regard to the fast decision making that was required in performing the RT task competence (as compared to the transfer task), competence is likely to be a matter of small degrees which we believe would be affected by our manipulation in the present context (if there were an actual effect).

How should we relate the present null findings to other positive findings in the literature (e.g., Han & Black, 2011; Kontra et al., 2015)? We believe the key difference is that the present study differs from previous studies on learning science concepts through physical interaction as we manipulated the lawful information that physical interaction affords, as opposed to contrasting different modes of physical interaction (e.g., mouse-based versus haptic manipulation; e.g., Skulmowski et al., 2016; Zacharia et al. (2012); no manipulation versus haptic manipulation, e.g., Han & Black, 2011; Kontra et al., 2015). As such, we aimed to exclude effects that can be attributed to different modes of physical interaction. Of course, this is not to say that previous studies that revealed an effect of different modes of physical interaction cannot be attributed to the lawful information that is afforded by these different modes of physical interaction. In fact, if embodied learning theories are correct, positive effects of physical interaction should be explained in terms of meaningful correspondences with the learning content (Pouw et al., 2014). Yet, if our interpretation of the current (unexpected) results is on track, caution is advised when attributing effects of physical interaction based on meaningful correspondences that may exist between action and science concepts. Therefore, future research could focus more on manipulating structural information that emerges out of perception and action loops (by loosening or tightening the correspondence between action and its perceptual correlates) rather than manipulating the perception–action loop altogether (i.e., manipulating the mode of interaction).

In sum, the current findings are interesting because they show that physical experiences in adults are not readily or easily integrated with the knowledge schemas of the kind that allows one to solve the performance tasks reported here. This resonates well with findings regarding physics misconceptions, which show that incorrect knowledge schemas are not easily altered by concrete counterevidence (Duit & Treagust, 2012). Thus future research could focus more on longer bodily training, and more specifically how this affords learners meaningful information (rather than mere physical engagement) that is *not* provided by the visual modality alone.

Acknowledgements—This research was funded by the Netherlands Organisation for Scientific Research (NWO-PROO, project number 411-10-908). We would like to thank Bonnie van Huik for her initial contributions to the preparation of the learning material. We would also like to thank two anonymous reviewers for their valuable comments on the original manuscript.

NOTES

1 For an overview of research on the role of Nintendo Wii-Board in learning and cognition, see, e.g. Dijkstra, Eerland, Zijlmans, and Post (2014) and Vernadakis, Gioftsidou, Antoniou, Ioannidis, amd Giannousi (2012).

2 Before the experiment the participants engaged in a calibration session: After a 3 s countdown, pressure data of sensors on both sides were recorded for the duration of 5 s at a sampling rate of 60 Hz. For each side of the Wii Board, the average recorded force on that side was subtracted from force values resulting in a new calibrated force value for each sensor (minus the weight of the hands in rest state). This ensured that the interface

only reacted when participants actively engaged with the Wii Board.

3 These analyses were performed on the complete sample of 90 participants. They did not include Wii-Board data of the training phase.

REFERENCES

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*, 637–660. doi:10.1017/S0140525X9953214

Bivall, P., Ainsworth, S., & Tibell, L. A. (2011). Do haptic representations help complex molecular learning? *Science Education*, *95*(4), 700–719. doi:10.1002/sce.20439.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes*, *27*, 594–610. doi:10.1080/01690965.2011.567074

Cutica, I., & Bucciarelli, M. (2013). Cognitive change in learning from text: Gesturing enhances the construction of the text mental model. *Journal of Cognitive Psychology*, *25*, 201–209. doi:10.1080/20445911.2012.743987

Cutica, I., Ianì, F., & Bucciarelli, M. (2014). Learning from text benefits from enactment. *Memory and Cognition*, *42*, 1026–1037. doi:10.3758/s13421-014-0417-y

Daponte, P., De Vito, L., Picariello, F., & Riccio, M. (2014). State of the art and future developments of the augmented reality for measurement applications. *Measurement*, *57*, 53–70. doi:10.1016/j.measurement.2014.07.009

Dijkstra, K., Eerland, A., Zijlmans, J., & Post, L. S. (2014). Embodied cognition, abstract concepts, and the benefits of new technology for implicit body manipulation. *Frontiers in Psychology*, *5*, 757. doi:10.3389/fpsyg.2014.00757

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*, 254–262. doi:10.1016/S1364-6613(98)01188-7

Duit, R., & Treagust, D. F. (2012). How can conceptual change contribute to theory and practice in science education? In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 107–118). Dordrecht, The Netherlands: Springer.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, *26*(1), 9–25. doi:10.1007/s10648-014-9249-3

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*, 424–436. doi:10.1037/0022-0663.96.3.424

Goldstone, R. L., & Barsalou, L. (1998). Reuniting perception and conception. *Cognition*, *65*, 231–262.

Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers and Education*, *57*, 2281–2290. doi:10.1016/j.compedu.2011.06.012

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86–104. doi:10.1037/a0034008

Jones, M. G., Minogue, J., Tretter, T. R., Negishi, A., & Taylor, R. (2006). Haptic augmentation of science instruction: Does touch matter? *Science Education*, *90*(1), 111–123. doi:10.1002/sce.20086

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. doi:10.1080/01621459.1995.10476572

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, *24*, 313–334. doi:10.1080/01690960802365567

Kiefer, M., & Trump, N. M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends in Neuroscience and Education*, *1*, 15–20. doi:10.1016/j.tine.2012.07.002

Klatzky, R. L., & Lederman, S. J. (2002). Touch. In I. B. Weiner, A. F. Healy, & R. W. Proctor (Eds.), *Handbook of psychology: Vol. 4. Experimental psychology* (pp. 147–176). New York, NY: Wiley.

Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical experience enhances science learning. *Psychological Science*, *26*, 737–749. doi:10.1177/0956797615569355

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic books.

Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education*, *8*, 74–88. doi:10.1111/mbe.12047

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, *5*, 196–211. doi:10.1111/j.1751-228X.2011.01129.x

Macedonia, M., & von Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biolinguistics*, *6*, 393–416.

Manches, A. D., & O'Malley, C. (2012). Tangibles for learning: A representational analysis of physical manipulation. *Personal and Ubiquitous Computing*, *16*, 405–419. doi:10.1007/s00779-011-0406-0

Marley, S. C., & Carbonneau, K. J. (2014). Theoretical perspectives and empirical evidence relevant to classroom instruction with manipulatives. *Educational Psychology Review*, *26*(1), 1–7. doi:10.1007/s10648-014-9257-3

Marshall P., Price, S., & Rogers, Y. (2003). Conceptualising tangibles to support learning. In *Proceedings of the 2003 conference on interaction design and children* (pp. 101–109). New York, NY: ACM Press.

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*, 679–690. doi:10.3758/s13428-010-0049-5

Nathan, M. J. (2012). Rethinking formalisms in formal education. *Educational Psychologist*, *47*(2), 125–148. doi:10.1080/00461520.2012.667063

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. doi:10.1016/0028-3932(71)90067-4

Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education*, *96*(1), 21–47. doi:10.1002/sce.20463

Pouw, W. T. J. L., Van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives.

Educational Psychology Review, *26*(1), 51–72. doi:10.1007/s10648-014-9255-5

Ryle, G. (1945). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society*, *46*, 1–16.

Schönborn, K. J., Bivall, P., & Tibell, L. A. (2011). Exploring relationships between students' interaction and learning with a haptic virtual biomolecular model. *Computers and Education*, *57*, 2095–2105. doi:10.1016/j.compedu.2011.05.013

Skulmowski, A., Pradel, S., Kühnert, T., Brunnett, G., & Rey, G. D. (2016). Embodied learning using a tangible user interface: The effects of haptic perception and selective pointing on a spatial learning task. *Computers and Education*, *92*, 64–75. doi:10.1016/j.compedu.2015.10.011

Triona, L. M., & Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction*, *21*(2), 149–173. doi:10.1207/S1532690XCI2102_02

Vernadakis, N., Gioftsidou, A., Antoniou, P., Ioannidis, D., & Giannousi, M. (2012). The impact of Nintendo Wii to physical education students' balance compared to the traditional approaches. *Computers and Education*, *59*, 196–205. doi:10.1016/j.compedu.2012.01.003

Wiebe, E. N., Minogue, J., Jones, G. M., Cowley, J., & Krebs, D. (2009). Haptic feedback and students' learning about levers: Unraveling the effect of simulated touch. *Computers and Education*, *53*, 667–676. doi:10.1016/j.compedu.2009.04.004

Zacharia, Z. C., & Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the physics by inquiry curriculum: The case of undergraduate students' conceptual under standing of heat and temperature. *American Journal of Physics*, *76*, 425–430. doi:10.1119/1.2885059

Zacharia, Z. C., Loizou, E., & Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Childhood Research Quarterly*, *27*, 447–457. doi:10.1016/j.ecresq.2012.02.004

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, *21*, 317–331. doi:10.1016/j.learninstruc.2010.03.001