Grounding Graph Theory in Embodied Concreteness with Virtual Reality

Julia Chatain*
Rudolf Varga*
ETH Zurich, Dept. of Computer Science, Dept. of Humanities, Social and Political Sciences
Switzerland

Violaine Fayolle
ETH Zurich, Dept. of Computer Science
Switzerland

Manu Kapur
ETH Zurich, Dept. of Humanities, Social and Political Sciences
Switzerland

Robert W. Sumner
ETH Zurich, Dept. of Computer Science
Switzerland

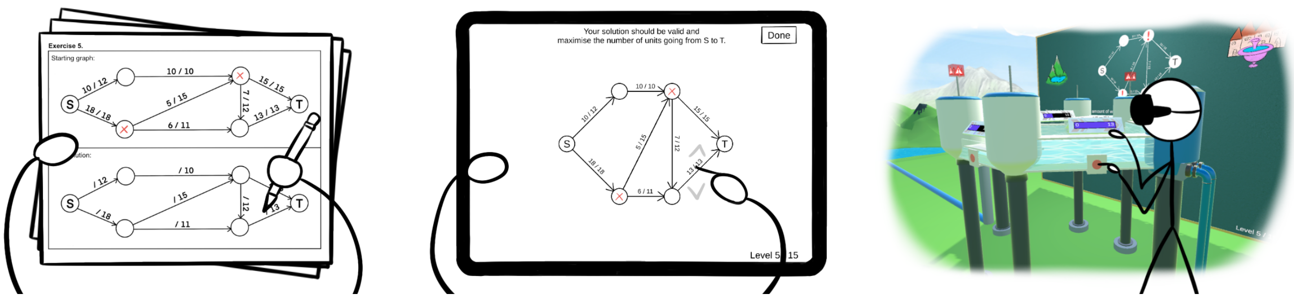

Figure 1: Three approaches to graph theory exercises: abstraction on paper (*ABST*), manipulated concreteness on tablet (*MNPL*), and embodied concreteness in Virtual Reality (*EMBD*).

ABSTRACT

Abstract mathematics can be difficult to grasp, in part because it relies on symbols and formalisms that are powerful yet meaningless to novices unless grounded in concreteness. Although a wide corpus of research focuses on concreteness in mathematics education, the notion of concreteness can be apprehended in various ways and it is not yet clear which specific aspects of concreteness help the learners. In this paper, we explore embodiment as a form of concreteness to ground abstract mathematics. First, we designed and evaluated an embodied learning activity on graph theory. Through a user study with 89 participants, we then compared three approaches: abstraction, manipulated concreteness, and embodied concreteness. Our results show that, compared to abstraction, both forms of concreteness increase learners' perceived attention, confidence, and satisfaction. However, only embodied concreteness increases perceived relevance and grounding. Moreover, unlike manipulated concreteness, embodied concreteness does not impair learning outcomes nor transfer abilities.

*Both authors contributed equally to this research.

TEI '23, February 26-March 1, 2023, Warsaw, Poland

ACM ISBN 978-1-4503-9977-7/23/02…$15.00
https://doi.org/10.1145/3569009.3572733

CCS CONCEPTS

• **Human-centered computing → Empirical studies in interaction design**; *User studies*; *Virtual reality*; **Gestural input**; **Empirical studies in HCI**.

KEYWORDS

embodied interaction, embodied cognition, embodiment, concreteness, mathematics education, virtual reality, problem-solving followed by instruction

ACM Reference Format:
Julia Chatain, Rudolf Varga, Violaine Fayolle, Manu Kapur, and Robert W. Sumner. 2023. Grounding Graph Theory in Embodied Concreteness with Virtual Reality. In *TEI '23: Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '23), February 26-March 1, 2023, Warsaw, Poland.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3569009.3572733

1 INTRODUCTION

Studies in math education indicate that many students hold unproductive beliefs about mathematics: they believe mathematics has nothing to do with the real world, and that one must learn solutions to math problems by heart [75]. For many of these students, mathematics is difficult to grasp. This is, in part, due to the fact that some of the most powerful aspects of mathematics rely on abstract symbols and formalisms that have no meaning, unless grounded in concreteness and provided with an interpretation [33, 34, 58, 91].

But when we talk about abstract mathematics, what are we really talking about? When a mathematician claims that she loves abstraction, while a student protests that he dislikes mathematics because "it is too abstract", are they really talking about the same thing? Similarly, when said student wishes mathematics were more concrete, is a concrete example truly the solution?

There is a verbal dispute in the field of concreteness for mathematics education. Although concreteness has been widely explored as a means to ground abstract mathematics, experts do not always align on their usage of the terms "concrete" and "abstract". For example, looking at only two papers in the field [28, 66], we found that the word "concrete" was associated with a wide range of terms: meaningful, familiar, well-understood, physical, grounded, pictural, perceptual, enactive, real-life, context-specific, and informal. In contrast, the word "abstract" was linked to: general, structural, portable, symbolic, vague, schematic.

What makes a good concrete example? Is it meaningfulness or physicality? Moreover, what is there to appreciate in abstraction? Is it generality or vagueness?

In this paper, we describe different kinds of concreteness, and offer an embodied perspective on the matter. Indeed, both embodied cognition and embodied interaction theories highlight the major role of users' bodies in meaning-making processes and grounding abstract concepts in the real world [21, 58, 79]. To illustrate the grounding capabilities of embodied concreteness in mathematics education, we designed and implemented an embodied activity to teach graph theory. We then used our activity in a user study to demonstrate the effect of different kinds of concreteness on motivation and learning outcomes. Specifically, we compared three approaches: abstraction, manipulated concreteness, and embodied concreteness.

Our paper illustrates the importance of rigorously defining concreteness and contributes with empirical evidence in favor of embodied concreteness for grounding abstract mathematics.

2 RELATED WORK

In this section, we discuss the role of concretenesses in learning abstract mathematics, and describe embodied approaches in this context.

2.1 Learning by grounding in concreteness

Learning abstract mathematics is difficult, in particular as, to novices, the related concepts and meaningless conventional symbolic systems can be difficult to grasp [33, 34]. One way of addressing this issue is by grounding mathematics in concreteness. Grounding is the process of mapping "novel ideas and symbols to modality-specific experiences that are personally meaningful" [58]. In particular, through grounding, a mapping is "formed between an idea or symbol, and a more concrete referent, such as an object, movement or event in the world - as well as mental re-enactment of these experiences - in service of meaning-making" [17, 58].

2.1.1 Concretenesses. Before going any further, we ought to define the term "concreteness". Indeed, although concreteness is often discussed in mathematics education, experts do not always align on their definition of concreteness. Identifying such verbal disputes is crucial as they can be tools for progress [11]. In this section, we highlight some of the main definitions of "concreteness" in the field of mathematics education.

First, an element can be thought as concrete if it can be touched, felt, smelt, kicked: if it can be *sensed* [93]. In that sense, a flower is more concrete than intelligence. This aspect can be influenced by technology, as certain elements can be made visible, for example light paths [26], or tangible, for example chemical forces [57]. This concreteness is also influenced by the learners' bodies, as they are central to sensory perception [43, 79].

Second, in fields such as mathematics and computer science, an element is often qualified as more concrete if it is more *specific*, constrained, precise, as opposed to general, overarching, and reusable [93]. For example, the sequence "{1, 2, 3}, List<int>, List<T>" evolves from more concrete to more abstract. This concreteness depends solely on the element itself and its context.

Finally, Dewey contrasts a concrete element that can be "readily apprehended by itself" to a more abstract one that can be "grasped only by first calling to mind more familiar things and then tracing out connections between them and what we do not understand" [18]. Here, an element is more concrete if it is more *relatable*, familiar, or imaginable, and abstract if it is unrelatable, unfamiliar, or unimaginable. With this definition, concreteness is not a property of the element alone, but rather a property of the element as perceived by the learner [93]. For example, Papert explains how a gears mechanism was a "comfortable friend" that helped him grasp the "otherwise abstract" concept of equations [63]. From this perspective, learning is a process of concretion, where one grows connections with an abstract element until it is relatable, and in that sense, concrete [93].

This distinction is important as not all concretenesses align. For example, the mathematical expression $2x - 4 = 0$ is abstract as it has no smell, and cannot be touched. However, for a mathematician, it can be very concrete as it is relatable and familiar. Working with this distinction is important to clearly identify the specific aspects of concreteness that are beneficial for grounding and, therefore, impact learning.

In this work, we consider different kinds of concreteness. Moreover, we do not consider concreteness and abstraction as categories separating elements into two sets, but as relative concepts discriminating elements along a spectrum. Therefore, in this work, the words "concrete" and "abstract" stand for "more concrete" and "more abstract". Moreover, in our concreteness study, we define our conditions from the most abstract condition, to which we add elements of concreteness. Therefore we refer to the most abstract condition as "abstract" and the other conditions as "concrete".

2.1.2 Grounding in concreteness. Several ways of grounding mathematics in concreteness have been explored, with mixed results. Kaminski et al. showed that using concrete, specific, relatable examples, as opposed to using abstract representations, is detrimental to transfer of knowledge [42], that is applying said knowledge to related yet different problems [65]. However, a replication of this study showed that the advantage of abstract representations disappeared when improving the concrete examples to make them more intuitive and less distracting [88].

Beyond visual representations, manipulable representations have also been explored, and show great potential in mathematics education [10]. In particular, the use of concrete manipulatives increases retention, problem solving, transfer, and justification scores over more abstract symbols [10].

When grounding mathematical concepts in concreteness, the process used is also of importance. For example, simultaneous multiple representations can help students ground abstract content to more visual artifacts [68]. However, several issues might arise. First, the representation dilemma: as students have to conjointly learn the novel content and the novel representation, one has to ensure that the benefits of the novel representation exceed its cost [68]. Second, using representations that are related but spatially or temporally distant can result in a negative split attention effect [84]. Another approach to grounding in concreteness is "concreteness fading", an instructional design building sequentially from a concrete, specific, relatable example to the corresponding abstract, general representation [53, 82]. Concreteness fading was proven beneficial over using solely concrete examples or abstract representations and over progressing from abstract to concrete representations [27, 28]. Traditionally, concreteness fading evolves from an enactive representation, to an iconic representation and concludes with a symbolic representation [82]. But other forms of concreteness could be explored as well, and, we believe, this field could also benefit from a clarification of the role of different aspects of concreteness in learning.

In this work, we explore the influence of different forms of concreteness on grounding, in the context of abstract mathematics. In particular, we focus on manipulated concreteness and embodied concreteness, described in the following section.

2.2 Learning through embodiment

Another way of grounding abstract mathematics is through embodiment. In this section, we define embodiment via embodied cognition and embodied interaction theories, and argue that embodiment can also be studied as a form of concreteness.

2.2.1 Embodied cognition. Embodied cognition theory stipulates that learners' bodies are involved in the learning process, either through action and manipulation, or as the primary constituent of cognition [2, 54]. From this perspective, thinking is described as a form of truncated action, where the sensorimotor processes related to the action are engaged, but not tangibly expressed externally [1]. Moreover, cognition is also situated: that is, the construction of knowledge happens through interaction with a temporal and physical environment [70].

Although mathematics is often though as disembodied, experts argue that the essence of mathematics arises from our embodied and situated relationship with our environment [50] and that discarding the embodied account is detrimental to mathematics education [1]. In conclusion, sense-making of abstract concepts is grounded in our embodied and situated relationship with the world [73].

Embodied approaches to mathematics learning have been explored, in non-digital as well as digital forms [87]. For example, directed bodily actions improved the quality of students' proofs in geometry [59]. Moreover, embodied approaches have been used to build on the need of children for sensory regulation by integrating fidgeting and balancing in the learning activity [86]. Furthermore, direct embodied and enacted approaches have been explored to teach derivatives, demonstrating the importance of aligning the design of the learning activity with the design of the interaction [13].

2.2.2 Embodied interaction. In turn, proponents of embodied interaction insist that interaction design should be tightly connected with the physical and social context of the interaction with digital content [21]. Instead of solely considering users' bodies as physical entities utilized to embed the interaction in a 3D context (*Körper*), the embodied interaction perspective suggests to focus on users' bodies as feeling entities (*Leib*) [56]. Moreover, approaches such as somaesthetic appreciation design describe how users' bodies can be integrated in the design process [35].

Embodied interactions have been investigated in playful as well as learning activities. Chatain et al. described the Digital Gloves mechanism, co-locating input and display, as a mean to bring users' bodies at the core of the interaction and reduce split attention effects [14]. Pei et al. explored hand gestures and mimes as novel hand interfaces to interact with digital content, for example by mimicking scissors to cut digital paper [64]. Focusing on physical context, Gervais et al. demonstrated how interaction can be expanded beyond the computer screen to the entire desk of the user [31].

Hereinafter, we use the more general term "embodiment" as "embodied interaction for embodied meaning-making". Embodiment can be implemented at different degrees, based on three constructs: sensorimotor engagement, gestural congruency, and immersion [41]. In this paper, we define "embodied concreteness" as a form of concreteness that involves a high degree of embodiment, in a situated and relatable context.

3 DESIGN

We started by designing and implementing an activity to ground graph theory in embodied concreteness. In this project, we focus on the max-flow problem [25, 72], where, given a graph, the student has to maximize the amount of flow traveling from the Source (S) to the Sink (T), while respecting the maximum capacity of the edges, and the fact that vertices cannot store units (Figure 2).

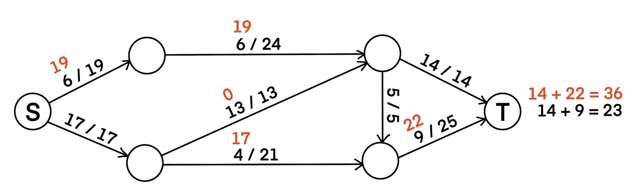

Figure 2: Example of a valid flow graph, as well as suggested modifications (in red, above) to maximize the flow value of the graph.

In this section, we describe how we designed, implemented, and validated our embodied concreteness activity.

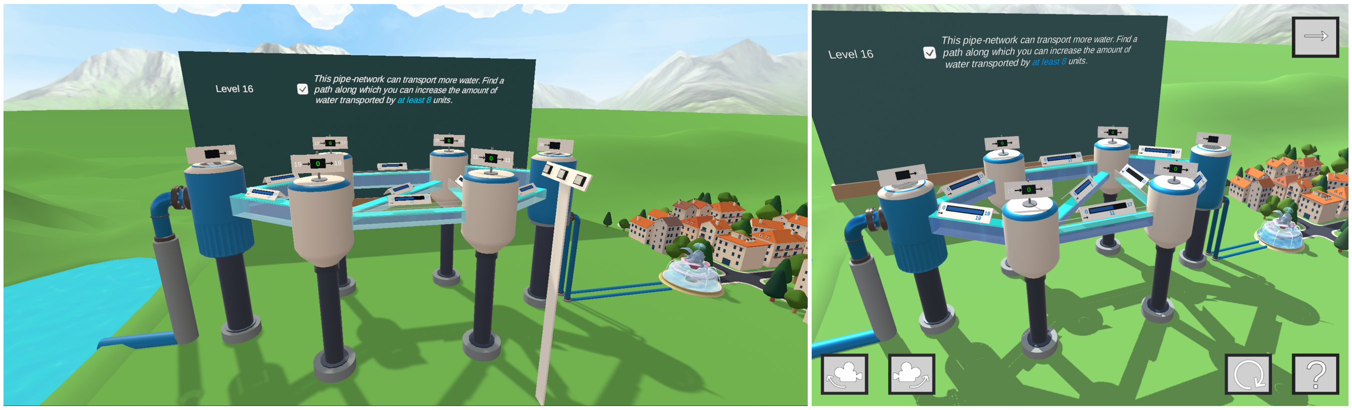

Figure 3: Initial prototype of the embodied graph theory activity in Virtual Reality (Left), as well as the Tablet implementation used as a control for the usability evaluation (Right). In the Virtual Reality condition, the water flow in an edge is manipulated by holding the bottom of the pipe with one hand, and moving the other hand up or down to indicate the desired water level. In the tablet condition, the level is adjusted by touching the bottom of the pipe with one finger, and adjusting the level by moving another finger up and down.

3.1 Concrete graph representation

Graphs can be represented in various ways, such as symbolically and geometrically. For our project, we focused on embodied concreteness, and therefore designed an embodied, sensed, situated, and relatable graph representation. To do so, we relied on embodied schemata from conceptual metaphor theory [49]. According to Lakoff and Johnson, "the essence of a metaphor is understanding and experiencing one kind of thing in terms of another". In addition, embodied schemata are "recurrent patterns of bodily experience". Specifically, we looked for bodily experiences that could be used as metaphors for flows in graphs.

To our knowledge, embodied schemata have not yet been explored in the context of graph theory. However, in the case of electricity and electrical networks, two main schemata are used: WATER-FLOW and MOVING-CROWD [30]. Reusing these schemata in the context of graph theory is particularly relevant as flow networks are often used to solve electrical networks problems [4, 15, 23].

We used the WATER-FLOW schema as it is the most commonly used and therefore most relatable (Figure 3, left). In our activity, a graph is represented as a pipe network (edges) between water towers (vertices), and a simple simulation of the water flow through the pipes is displayed in real time. The goal for the student was to increase or maximize the amount of water flowing from the lake (source) to the city (sink). Upon success, a fountain placed at the entrance of the city starts pouring water.

3.2 Embodied interaction with a graph

As to attain a high degree of embodiment, we implemented the activity in Virtual Reality (VR) [41], using hand tracking over controllers. Although interaction with graphs in VR has already been explored, previous work focused more on data visualization and manipulation, and thus did not fit out project [22, 39]. Therefore, we designed our own system.

To design the embodied interaction with the edges of the graph, we looked at two approaches: direct-embodied and enacted [54, 62]. The type of embodiment influences the mathematical concept emphasized, and impacts learning outcomes and persistence [13]. As our activity focuses on the value of the flow in the edge, rather than its variation, we selected a direct-embodied approach.

During the development of the interaction, we evaluated various input mechanisms on a small group ($n = 6$) of participants who did not have a lot of experience with VR systems, in an informal setting. We gave the participants a short explanation on the input mechanism at the beginning and let them solve problems on their own. They gave verbal feedback on their experience, while we monitored their in-game activity. We wrote down the problems the participants faced during their time in the activity. After the use of the activity the participants explained their issues, thoughts and ideas in their own words regarding both the activity itself and the input technique. For the first participant, we used a one-handed input approach, setting the water level. Due to the imprecision when locking the water level, for further participants we introduced a two handed variant separating the adjustment and the locking movements. The locking was deactivated by removing the second hand from the pipe. We found that the participants were still struggling as they removed both hands at the same time. As a result, the interaction goes as follows: with one hand, the learner grabs the bottom of the edge with an open-close gesture, and by moving the other hand vertically, they can adjust the amount of water flowing in the edge. This two-handed interaction has two main advantages: it strengthens the embodiment as the capacity of the edge is directly congruent to the distance between the two hands, and it improves usability as the learner can release the lower hand to precisely set the level of the edge.

To improve usability and embodiment, we also adjusted the height of the entire graph to fit the height of the learner.

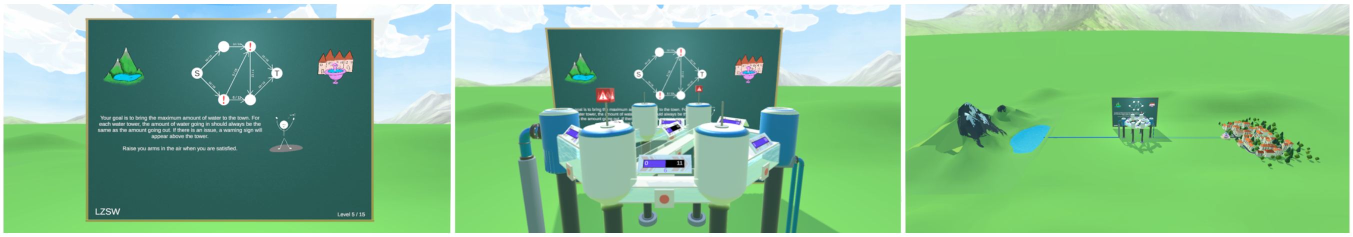

Figure 4: Final prototype of the embodied graph theory activity in Virtual Reality. In this prototype, the water flow in an edge is manipulated by pressing the button at the bottom of the pipe with one hand, and moving the other hand up or down to indicate the desired water level.

3.3 Pedagogical pattern

In our work, we focus on grounding, and therefore we needed to reconnect our activity with more formal forms of instruction [13]. To do so, we decided to follow a Problem Solving followed by Instruction (PS-I) pedagogical pattern, where, as opposed to the more wide-spread I-PS pattern, the students solve exercises about the topic before receiving instruction [51]. This pattern was proven effective for mathematics learning and relies on three mechanisms: activation of prior knowledge, awareness of knowledge gap, and recognition of deep features [51, 78]. In our activity, we activate concrete prior knowledge by offering a relatable experience to the students. Moreover, we designed our levels to increase knowledge gap awareness and identification of deep features. Indeed, each level either increases complexity to encourage search for more general solutions, or highlights new problem features such as the counter-intuitive need to decrease flow on certain edges to increase output flow. The Instruction phase of the PS-I pattern was handled differently for each study and is described in their respective sections. We selected a PS-I pedagogical pattern over a concreteness fading one as we were interested in the effect of different forms of concreteness on grounding and learning. Choosing to focus on concreteness fading would imply defining a different sequential pattern for each condition and shift the focus away from our research question.

3.4 Design Validation

We validated our design with a user study focused on usability. To do so, we designed a control condition on a tablet (Figure 3, right). In this condition, the representation is the same, but the embodiment is of a lower degree as immersion and sensorimotor engagement are reduced [41]. For the tablet prototype, we replaced the two-hands interaction by a two-fingers interaction where one presses an edge with one finger, and adjusts its flow quantity with another finger. To mimic the navigation of the VR condition, we added two buttons to rotate the camera around the pipe network.

The goal of this study is to validate our VR prototype, in particular in terms of usability, and identify potential directions for improvement. Specifically, we want to ensure that technology does not impact the experience by reducing usability or increasing simulator sickness. We used a tablet version of the same activity as a control condition as most learners are used to touch screens, but not to VR.

3.4.1 Demographics. We recruited $n = 26$ participants (6 identifying as female, 20 as male), from Zurich, Switzerland ($n = 9$), and Budapest, Hungary ($n = 17$). Participants were, in average, $M = 27.26$ years old ($SD = 8.17$) and were assigned to the tablet ($n = 13$) and VR ($n = 13$) conditions randomly. None of the participants had VR experience while 18 participants had tablet experience. One participant in the tablet condition was removed from the analysis as she was an outlier in terms of time spent in the activity.

3.4.2 Protocol. We tested our prototypes within a PS-I pedagogical pattern. First, participants completed a general questionnaire including demographics questions, followed by a learning pre-test, and a Simulator Sickness Questionnaire (SSQ) [45]. Then, as a Problem Solving phase, the participants solved the graph theory problems with either the tablet prototype or the VR prototype. Afterwards, as an Instruction phase, the participants watched a short video on the Ford-Fulkerson algorithm [25, 72]. Finally, participants completed a System Usability Scale (SUS) questionnaire [8], a SSQ, as well as a learning post-test comprised of recall and transfer questions with different representations. This study was approved by the ETH Ethics Commission as proposal EK 2022-N-64.

3.4.3 Results. Aligned with previous work [13], we found no significant differences in usability between the tablet and the VR prototypes ($p = 0.14$, $t(24.0) = 1.52$). The tablet prototype received a SUS score of 86.54 ($SD = 6.89$), qualified as "Excellent" [5], while the VR prototype received a score of 81.35 ($SD = 10.19$), qualified as between "Good" and "Excellent".

We did not find significant differences in SSQ scores either ($p = 0.33$, $t(24.0) = -0.97$). In particular, the tablet prototype can be categorized as generating "negligeable" symptoms, while the VR condition generates "minimal" symptoms [80]. As our VR activity does not include fast-paced changes, simulator sickness is reduced [81].

The VR activity took more time that the tablet activity ($p = 0.06$, $t(24.0) = -2.00$). This is also congruent with previous research [13], and is justified by the fact that VR participants perform wider movements and moved around the space more.

3.5 Design improvements

Based on our observations during the study, we made several adjustments to our prototype (Figure 4). To improve embodiment and acknowledge the diversity of learners' bodies [79], we added a skin color selection panel. To make the learners more confident in their

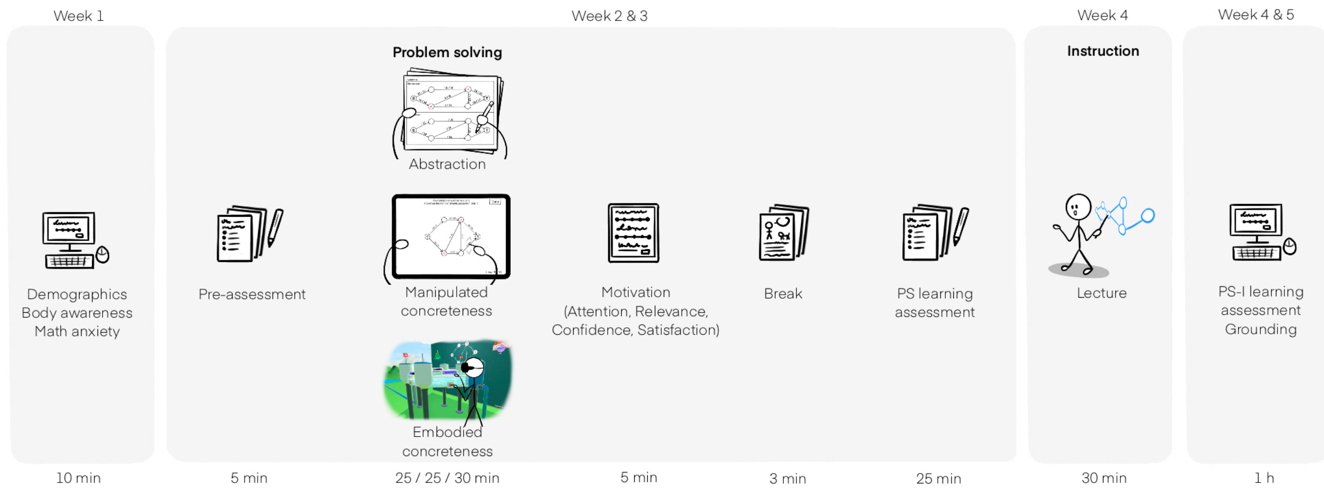

Figure 5: Overview of the user study protocol.

movements, we added a tutorial where they can explore the virtual space [13]. To improve usability, we made the direction of the edges clearer. As several participants reported struggling getting an overview of the problem, we added a depiction of the graph on the black board. We also added a button on the edges to make the interaction technique clearer and strengthen the embodiment (Figure 4, center). Finally, following the idea of "experiencing the body as play", we designed for a sense of embodied achievement at the end of each level by having the users adopt a "winning position", that is raising both arms in the air, to launch the next level (*Leib*) [56].

4 COMPARISON OF CONCRETENESSES

4.1 Research Questions

After validating our design, we used our activity to address the following research questions:

RQ1 What is the impact of concreteness on grounding?
RQ2 What is the impact of concreteness on learning outcomes?

In this work, we want to evaluate the impact of different forms of concreteness on grounding and learning. To do so, we focus on concreteness as a standalone Problem-Solving intervention in a PS-I pedagogical pattern. For this study, we designed three experimental conditions (Figure 1). In the abstraction condition (*ABST*), the students solve the exercises on paper, with a geometrical graph representation. This condition is the most abstract as it is not manipulated, embodied, situated, nor relatable for graph theory novices. In the manipulated concreteness condition (*MNPL*), the students solved the same exercises on a tablet, where they can interact with the graphs' geometrical representation. This condition is concrete as it is manipulated and embodied at a low degree. However, it is still not situated nor relatable. In the last condition, embodied concreteness (*EMBD*), the students solved the exercises using our embodied graph theory activity, in a highly embodied, situated, and relatable manner.

4.2 Demographics

We recruited $n = 89$ (33 female, 54 male, 0 other, 2 undisclosed) volunteer bachelor students, aged $M = 20.6$ years old ($SD = 2.00$), from a Data Structures and Algorithms course at the mathematics department of ETH Zurich. The participation was rewarded by gaining access to a bonus exercise awarding extra points to the final exam. Participants were randomly assigned to each of the conditions ($n_{ABST} = 30$, $n_{MNPL} = 29$, $n_{EMBD} = 30$). The intervention included a pre-assessment questionnaire to evaluate previous knowledge on graph theory and the specific max-flow problem. This assessment was designed by the authors and included items such as "Have you learned graph theory previously?" with answers such as "I have, in a formal environment. (e.g. secondary school or university)", "I have, only informally. (e.g. self-study)", "No, I have not.". The assessment also included a max-flow problem to solve. Only 3 participants reported having learned about problems involving graphs before, and only one of them managed to solve the max-flow problem successfully. This participant was excluded from the analysis and is not included in the $n = 89$ sample size.

4.3 Protocol

We used a between-participants design to avoid learning effects across conditions. The study was composed of five steps: pre-intervention, intervention (Problem Solving phase), lecture (Instruction phase), exercises and post-intervention questionnaire (Figure 5).

During the pre-intervention, participants completed a questionnaire at home, including general demographics questions, a body awareness questionnaire [77], and a math anxiety questionnaire [36].

The Problem-Solving intervention was conducted in our lab, in a separate room. For the *EMBD* condition, we prepared a VR space of 4m * 4.5m to accommodate all the levels of the activity. During the intervention, the participants signed a consent form and completed a pre-assessment asking about their knowledge of Graph Theory. They then solved max-flow problems in one of the three conditions: abstraction (*ABST*), manipulated concreteness (*MNPL*),

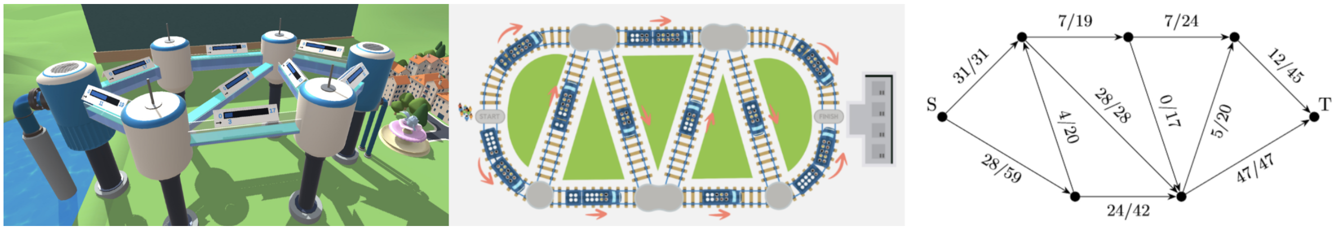

Figure 6: Post-test representations: concrete WATER-FLOW embodied schema, concrete MOVING-CROWD embodied schema, abstract.

or embodied concreteness (*EMBD*). Participants had 25 minutes to solve the problems, except *EMBD* participants who had 30 minutes in order to account for the calibration steps. For the *MNPL* and *EMBD* conditions, we logged the actions of the user. Then, participants filled in an Instructional Materials Motivation Survey following the ARCS model: Attention, Relevance, Confidence, and Satisfaction [44, 48, 52]. In order to alleviate fatigue effects, participants then took a three minutes break where they could read some selected comics. Finally, participants solved a 25 minutes learning assessment, evaluating the effect of the problem solving phase, and focusing on isomorphic problem-solving with different representations (Figure 6): concrete based on the WATER-FLOW embodied schema (similar to *EMBD* condition), concrete based on the MOVING-CROWD embodied schema (the graphs are represented as a train network with a flow of passengers), and abstract (similar to *ABST* and *MNPL* conditions) [30]. The questions were presented in a randomized order to alleviate effects undesirable within learning assessment, such as concreteness fading [9, 28].

About a week after the intervention, as an Instruction phase, participants followed a lecture on graph theory and max-flow problems, by their regular instructor, at their regular schedule.

A few days later, participants solved exercises on maximum flows on their usual exercise platform, and completed a questionnaire about the relevance and usefulness of the intervention for the lecture and the exercises.

This study was approved by the ETH Ethics Commission as proposal EK 2022-N-40.

4.4 Results

In the following sections, we performed one-way ANOVAs with the following contrasts: abstraction opposed to concreteness, and manipulated concreteness opposed to embodied concreteness. We checked for the assumption of normality with a Shapiro-Wilk Normality Test [71], and we checked for the assumption of homoscedasticity using a Breusch-Pagan Test [7]. If the assumptions were met, we used a regular ANOVA [12] with Bonferroni post-hoc comparisons, otherwise we used a robust ANOVA [92] with Linear Constraints post-hoc comparisons.

To evaluate learning outcomes, we performed a Bayesian analysis to make sense of non-significant statistical tests [19, 24, 40, 89]. To do so, we first performed a Bayesian two-sided analysis of variance, followed by post-hoc tests when necessary. Specifically, we considered equal prior odds and compared a null model (M_0) to a model considering the main effect of condition only (M_1). In the following, we use BF_{01} to describe the ratio $\frac{P(M_0)}{P(M_1)}$.

4.4.1 Grounding. To address RQ1, we first looked at the four components of the ARCS model [44]. Attention refers to how captivating and interesting the content is for the learners. Relevance refers to how valuable and connected to the real world the experience is. Confidence refers to how much the activity helped the learners feel in control of their success and likely to succeed. Satisfaction refers to how good the learners feel about their accomplishments and continuing to learn.

We found significant effects of condition on all four components (Figure 7): Attention ($F(2, 34.64) = 16.06$, $p < 0.001$), Relevance ($F(2, 86) = 21.72$, $p < 0.001$), Confidence ($F(2, 34.28) = 4.86$, $p = 0.014$), and Satisfaction ($F(2, 33.54) = 5.46$, $p = 0.009$). The effect sizes were large for Attention ($\xi = 0.64$) and Relevance ($\eta = 0.58$), and medium for Confidence ($\xi = 0.40$) and Satisfaction ($\xi = 0.45$).

Regarding Attention, there was a significant difference between *ABST* and *MNPL* ($p = 0.001$), as well as between *ABST* and *EMBD* ($p < 0.001$), but not between *MNPL* and *EMBD* ($p = 0.15$). Similarly, regarding Confidence, there was a significant difference between *ABST* and *MNPL* ($p = 0.034$), as well as between *ABST* and *EMBD* ($p = 0.028$), but not between *MNPL* and *EMBD* ($p = 0.81$). Regarding Satisfaction, we found a close to significant difference between *ABST* and *MNPL* ($p = 0.089$), and significant difference between *ABST* and *EMBD* ($p = 0.007$), but there was no significant difference between *MNPL* and *EMBD* ($p = 0.22$). In contrast, for the Relevance component, we found no significant difference between *ABST* and *MNPL* ($p = 0.41$), but we found a significant difference between *ABST* and *EMBD* ($p < 0.001$) as well as between *MNPL* and *EMBD* ($p < 0.001$).

We also looked at the results of the grounding questionnaire (Figure 8). This questionnaire included 5 points Likert scale items such as "Did the activity make you excited about joining the lecture?" or "Do you feel that the activity prepared you for the lecture?". Similar items about solving the final exercises were included.

We found a significant effect of condition on the grounding items related to the lecture ($F(2, 28.53) = 3.87$, $p = 0.033$), with a medium effect size ($\xi = 0.38$). The post-hoc analysis only revealed significant difference between the *ABST* and the *EMBD* conditions ($p = 0.030$), while no significant difference was revealed between the *ABST* and *MNPL* conditions ($p = 0.23$), nor between the *MNPL* and *EMBD* conditions ($p = 0.23$).

We did not find a significant effect for the items related to the exercises ($F(2, 75) = 0.60$, $p = 0.55$), but, although many participants

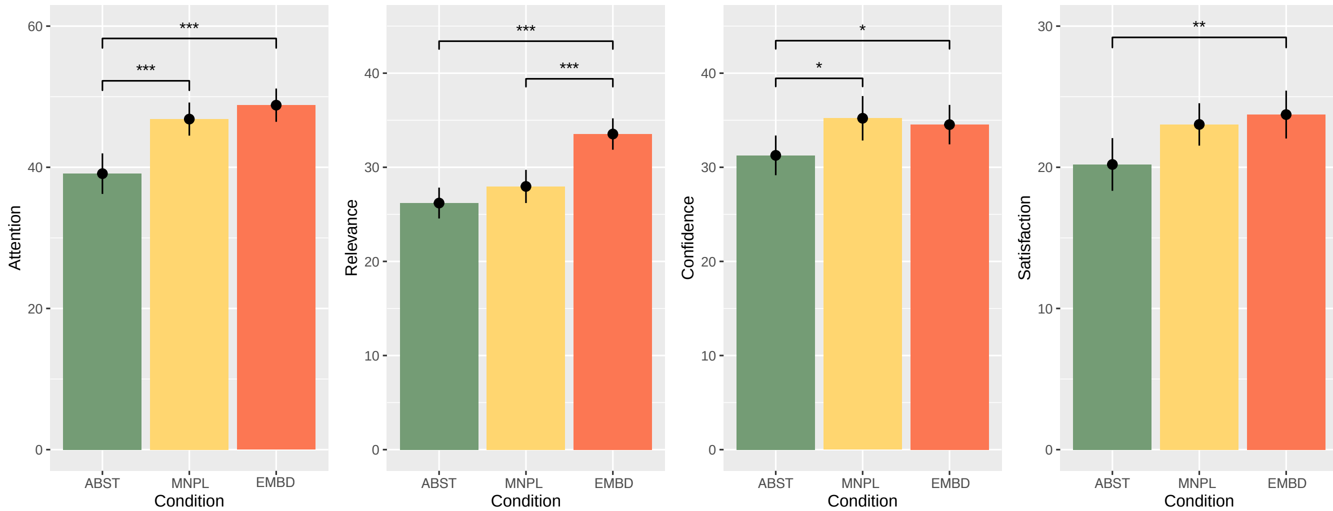

Figure 7: Bar plot representation of the ARCS model per condition (abstraction, manipulated concreteness, embodied concreteness), with adjusted p-values (*p < 0.05, **p < 0.01, ***p < 0.001).

answered the questionnaire, too few participants actually solved the exercises. Therefore, we refrain from drawing any conclusions regarding this aspect.

In conclusion, regarding RQ1, both forms of concreteness significantly improved Attention, Confidence, and Satisfaction for the learners. However, only embodied concreteness improved perceived Relevance. Moreover, only embodied concreteness improved perceived grounding after the lecture. This is particularly important as students often believe that mathematics has nothing to do with the real world [75]. Our results show that this issue can be alleviated using embodied concreteness.

4.4.2 Learning Outcomes. To address RQ2, we focused on learning outcomes. This is particularly important as previous research shows that the use of concrete examples may reduce the transfer capabilities of the students [42], and that, according to cognitive load theory [83], high-immersive gaming environments such as VR might impair learning outcomes [47]. As too few participants completed the PS-I learning assessments, we did not include this test in our analysis.

Following the aforementioned procedure, we performed ANOVAs with respective subsequent comparisons on the different learning assessments (Figures 6, 9). We found no significant effect of condition in the following learning assessments: WATER-FLOW representation ($F(2, 34.65) = 0.10$, $p = 0.90$) and MOVING-CROWD representation ($F(2, 34.06) = 0.48$, $p = 0.62$). The only significant effect was on the abstract representation ($F(2, 34.3) = 3.40$, $p = 0.045$), and it was of medium size ($\xi = 0.33$).

As absence of evidence is not evidence of absence, we performed a Bayesian analysis to estimate whether there is indeed no performance difference between the groups on the WATER-FLOW and MOVING-CROWD learning assessments. Our analysis revealed moderate evidence for the null hypothesis of no effect on the WATER-FLOW assessment ($BF_{01} = 9.706$) as well as the MOVING-CROWD assessment ($BF_{01} = 6.834$).

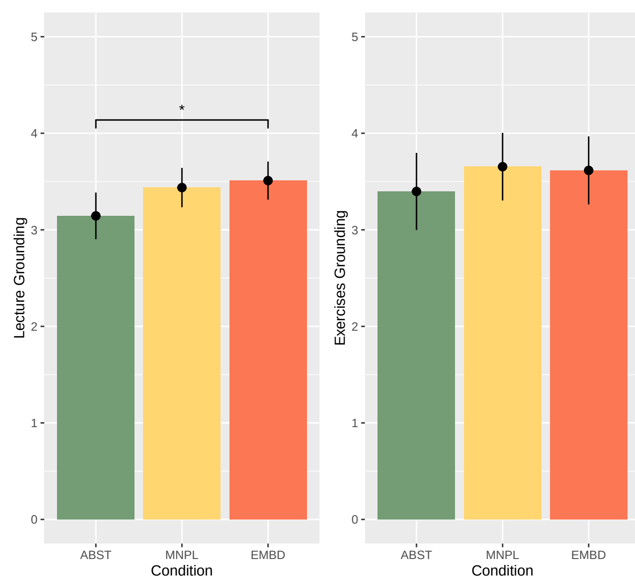

Figure 8: Bar plot representation of the perceived grounding after the lecture and the exercises, with adjusted p-values (*p < 0.05).

On the abstract representation exercises, students in the *ABST* condition ($M = 71.7$, $SD = 33.5$) outperformed students in the *MNPL* condition ($M = 53.81$, $SD = 31.5$) significantly ($p = 0.04$). However, there was no significant difference between the *ABST* condition and the *EMBD* condition ($M = 62.45$, $SD = 36.5$, $p = 0.34$), with anecdotal evidence for the null hypothesis ($BF_{01} = 1.982$). The post-hoc tests revealed anecdotal evidence of no effect between *ABST* and *EMBD* ($BF_{01} = 2.457$), as well as between *MNPL* and *EMBD*

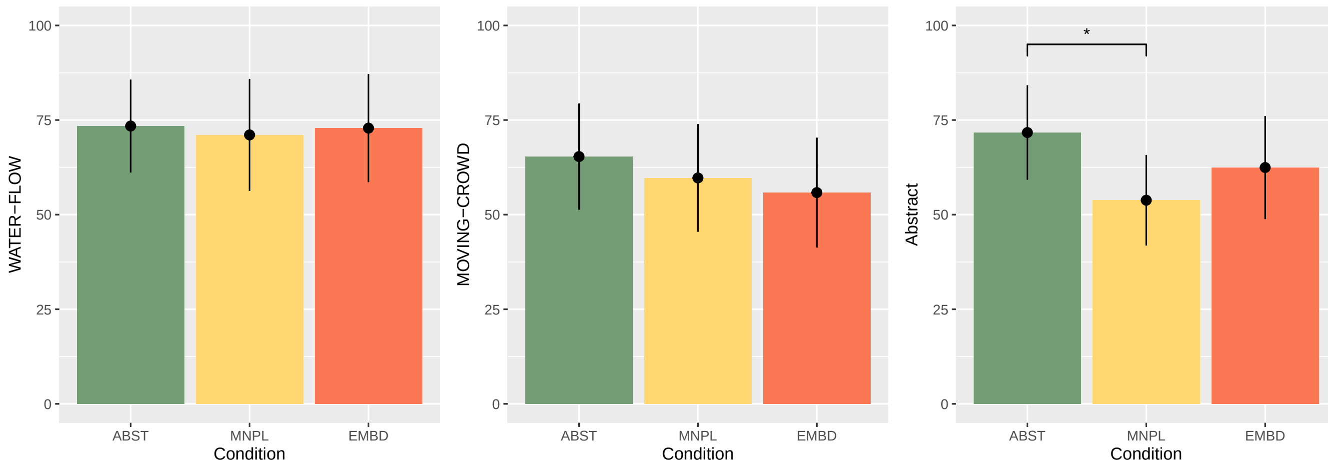

Figure 9: Bar plot representation of the learning outcomes for different representations, with adjusted p-values (*p < 0.05)

($BF_{01} = 2.548$). However, we observed anecdotal evidence for the alternative hypothesis between *ABST* and *MNPL* ($BF_{01} = 0.603$).

In conclusion, there was no effect of the condition on learning outcomes with concrete representations. However, students learning with manipulated concreteness performed worse on abstract tasks, while this difference was not significant for students learning with embodied concreteness. This means that embodied concreteness did not impair the capabilities of the students to transfer to a different embodied schema. It seems that embodied concreteness also did not impair transfer to a more abstract representation, although the evidence is only anecdotal at this point. However, manipulated concreteness had a negative impact on transfer, possibly as students relied too much on feedback, preventing them from thinking deeper about the problem.

5 DISCUSSION

In our work, we described several perspectives on concreteness and abstraction, and argued that there is a verbal dispute in the field on mathematics education. While many experts explore the role of concreteness in mathematics learning and teaching, the terms "concrete" and "abstract" are often under-specified, resulting in mixed results and ambiguity. We then argued that embodiment, that is embodied interaction for embodied meaning-making, can be explored as a form of concreteness.

To illustrate our discussion, we designed and implemented an embodied activity targeting the max-flow problem in graph theory. We then compared three approaches to problem-solving: abstraction, manipulated concreteness, and embodied concreteness. Our results demonstrate that different aspects of concreteness have different impacts on grounding and learning outcomes.

In the rest of the section, we discuss the mechanisms of learning with concreteness, the impact of our work, as well as its limitations, and provide suggestions for future work.

5.1 Mechanisms of learning with concreteness

Following on the outcomes of our user study, we discuss what mechanisms are responsible for these results, and in particular, which affective and cognitive mechanisms are activated by different kinds of concreteness. In the following, we only focus on the mechanisms involved in the problem solving phase, as looking at the entire PS-I pattern would be beyond the scope of our project, and has already been explored in previous work [78].

In our project, we explored two forms of concreteness. In the manipulated concreteness, students could manipulate a graph representation, with a low degree of embodiment. The system would give them limited feedback, for example prevent them from exceeding the capacity of an edge, or indicate when a node is invalid. In the embodied concreteness condition, the students could manipulate a situated and relatable representation of a graph, with a high degree of embodiment. The provided feedback included more information, as a water flow was also simulated along the pipes composing the graph.

As a result, different mechanisms should be considered for each of these conditions, summarized in Table 1. First, there are several mechanisms related to feedback only. Indeed, feedback supports error identification [55] and strategy acquisition [29]. Moreover, while the effect of low-information feedback is usually low, high-information feedback has a stronger impact as it supports error understanding [94].

Second, from a representation-agnostic standpoint, embodiment involves three main mechanisms [46]. Direct state induction is a mechanism of embodiment relying on the fact that certain bodily states impact the feelings of the learner independently of any cognitive mechanism. We support this mechanism in our embodied activity as we designed the experience from the feeling body (*Leib*) perspective [56], for example by inducing a feeling of embodied achievement as the learners adopt a winning position to finish a level. In turn, modal priming is a mechanism through which sensorimotor states enable learners to access abstract concepts, for example via conceptual metaphors. In our project, this mechanism is activated through the WATER-FLOW embodied schema [30, 49]. Finally, sensorimotor simulation is a mechanism of embodiment through which congruent bodily states and actions ease subsequent mental simulations, in particular in the context of problem

Table 1: The mechanisms of learning with concreteness.

Condition	Concreteness	Mechanisms
Abstraction		
Manipulated concreteness	Low-information feedback	Error identification Strategy acquisition
	Embodiment (low degree)	Sensorimotor simulation (low degree)
Embodied concreteness	High-information feedback	Error identification Error understanding Strategy acquisition
	Embodiment (high degree)	Direct state induction Modal priming Sensorimotor simulation (high degree)

solving [20]. We reconnect this particular mechanism to the conceptualization of thinking as truncated action [1]. In this new light, sensorimotor simulation describes how sensorimotor experiences can support further truncated actions, and therefore, thinking.

In our study, we showed that both forms of concreteness increased Attention, Confidence, and Satisfaction similarly. Therefore, these results should be explained by the common mechanisms between the two conditions: error identification, strategy acquisition, and sensorimotor simulation.

In turn, only embodied concreteness increases perceived relevance and grounding. We believe that this is explained by the modal priming mechanism as it reconnects the content to the learners' personal experiences, which is an important aspect of relevance in learning [76]. Moreover, relevance can be defined as a continuum of personal association, personal usefulness, and identification, and can trigger different mechanisms based on personal differences [67, 76]. In future work, such mechanisms should be explored in more depth in order to provide a more detailed account of the mechanisms of embodied concreteness.

Finally, learning with manipulated concreteness reduced the learning outcomes on abstract representations. We believe that this is explained by the lack of error understanding mechanism in this condition. In particular, as the representation was familiar to the students, they felt confident about solving the problems, and therefore, we believe, were more prone to errors.

5.2 Impact

With this work, we hope to impact the field of concreteness in mathematics in two ways.

First, we illustrated the need for a more rigorous definition of "abstraction" and "concreteness" in the field of mathematics education. In future work, we believe that a taxonomy of concretenesses should be defined, for example, building on a categorization framework of different representations along aspects of groundedness and idealization [6]. Moreover, we saw that concreteness can be defined as a property of the object only (concrete as specific), but also through the interaction of a learner with the object (concrete as tangible), or the mental model the learner has of the object (concrete as relatable). This aspect could be deepened if reconnected to the theory of affordances, building on the similar distinction between the Gibsonian and the Normanian perspectives [32, 61]. Such tool should then be used to support a meta-analysis of previous work on concreteness and mathematics education, and identify which aspects of concreteness, and related affective and cognitive learning mechanisms, specifically impact learning. Furthermore, investigating abstraction is at least as important, as the link between concreteness and abstraction is not necessarily dual, and similar verbal dispute exists for abstraction. For example, although abstraction is often conceived as a Platonic overarching, perfect ideal or truth, more recent work on abstraction offers an alternative grounded in mathematics history. For example, according to Wagner, mathematical abstraction can be defined as "the practice of incomplete, underdetermined, intermittent and open-ended translations between systems of presentations" [90], a horizontal paradigm often forgotten in mathematics education [3]. Moreover, vagueness, another word often associated with abstraction, can actually be formalized within the mathematical framework as vague or fuzzy mathematics [85]. Exploring different forms of abstraction would be particularly impactful within the field of concreteness fading [53].

Second, we showed that, although different kinds of concreteness can improve learners' attention, confidence, and satisfaction, embodied concreteness is a uniquely powerful tool for grounding mathematics as it increases perceived relevance while not impairing learning outcomes and transfer to more abstract representations. With embodied concreteness, learners can connect abstract concepts to real world experiences, thus challenging their unproductive beliefs about mathematics [75]. In future work, other comparisons should be explored. For example, comparing relatable but disembodied opposed to relatable and embodied would help isolate the effect of the modal priming mechanism in embodiment.

5.3 Limitations and Future Work

The main limitation of our work is the use or different technologies for the different conditions (paper, tablet, and VR). This was a conscious decision as we found important to offer a fair comparison by selecting the most appropriate technology for the activities we wanted to design. Indeed, offering non-manipulable graphs in VR, a technology heavily focused on bodily manipulations, would create unnecessary fatigue and confusion for the users. However, our solution is not perfect: different technologies come with different effects, that are non-negligible, such as novelty effect in the case of VR [38]. We believe this issue is mitigated as our usability study shows that there is no significant difference between our activity in VR and on tablet. However, to complement this work, further studies should investigate the role of technology in these results.

Another difference between our condition is the use of an overview graph in the embodied concreteness condition. This overview was added as our preliminary study revealed a perspective issue in this condition: on the tablet, the learner could see all the graph, while in VR, the learner had to navigate the space to build a global understanding of the problem. Another solution would have been to reduce the size of the network altogether. While we explored this solution through informal testing, we noticed that the lack of precision due to the hand tracking made the experience frustrating to the users. This might impact the learning outcomes, either by generating a negative split attention effect [84], or by inducing a positive indexing effect [37].

To address these limitations, the design of the second study could have been improved by adding another condition using the relatable pipe system representation, but on tablet, similar to the tablet condition of our usability study. This way, we would have been able to isolate the effect of manipulation from the effect of representation. However, in light of the context of our study, we only had access to a limited number of participants. Based on our power analysis, we could not afford reducing the number of participants per condition by adding another condition. We selected these specific conditions for our learning study for two reasons: First, we wanted to evaluate the potential of embodied concreteness for grounding, second, we wanted to isolate the effect of interaction and feedback as compared to a paper condition. However, to complete this contribution, future work should explore the role of representation decoupled from interaction.

Several other aspects could be improved in future work. First, the Problem Solving part of our intervention was conducted in our lab, which is not an ecologically valid environment. Second, in our study we only looked into short-term learning outcomes. Measuring learning outcomes over several months might reveal differences, for example, the effect of grounding on long-term learning outcomes. Moreover, we believe that including embodied assessments in the study design might reveal interesting insights. Indeed, learners are usually able to express understanding through gestures before they can articulate it with speech [16, 60, 69], and embodied assessments would capture this effect. Finally, the assessments were limited in time, and in English. This could have biased some of the results, in particular for slower students and non-native English speakers.

Another concern is the diversity of our sample. For example, only few women participated (23% in the first study, 37% in the second one). This is mostly due to our recruitment. For example, for the second study, we wanted to focus on participants from our target group, we recruited students' from the Mathematics department, already suffering from a gender diversity issue (23% women). Moreover, we only tested our approach with students from a mathematics Bachelor program. However, mathematics is a field of importance, even to those who do not wish to become mathematicians. Future work should focus on performing a similar study with a population less intrinsically motivated by the field. Another important concern to raise is that our sample only included able-bodied participants. However, a wider diversity of bodies should be included in embodiment research [79].

Finally, previous work identified the need to reconnect concreteness with more abstract representations [13, 27, 28]. In our studies, we did so using videos or lectures, but building this connection directly in the embodied activity might facilitate transfer. Moreover, we focused on concreteness in a standalone intervention, and we did not explore the impact of these different forms of concreteness within a sequential pattern. Future work should investigate embodied concreteness fading, for example using tools for embodied input of mathematical expressions [74].

6 CONCLUSION

In our paper, we explained that students often struggle to grasp mathematics as it heavily relies on abstract symbols and formalisms that only gain meaning when grounded in concreteness. We then highlighted a verbal dispute in the field of concreteness in mathematics education, as the word "concrete" is not always used with the same meaning. In particular, we presented different kinds of concreteness and described embodiment as a powerful form of concreteness as embodied experiences support meaning-making through interaction with relatable objects and environments.

To support our argument, we created an activity to solve graph theory problems in an embodied manner, and validated our design with a first user study, revealing its high usability. We then used our activity to demonstrate the effect of different kinds of concreteness. Through this second user study, we compared three conditions: abstraction, manipulated concreteness, and embodied concreteness. Our results show that both forms of concreteness can increase learners' attention, confidence, and satisfaction. However, only embodied concreteness increases perceived relevance and supports grounding. Moreover, unlike manipulated concreteness, embodied concreteness did not negatively impact performance on abstract representations.

With this work, we contribute to the field of mathematics education in two ways. First, we illustrate the importance of rigorously distinguishing different kinds of concreteness. Second, we provide empirical evidence supporting embodied concreteness as a powerful tool to ground abstract mathematics.

ACKNOWLEDGMENTS

The authors would like to thank Dražen Popović, Keny Chatain, Anna Ettlin, Charlotte Müller, Vera Baumgartner, Samuel Tobler, GTC and LSE for the support and feedback, Édouard Angebault and Fraser Rothnie for some of the graphical elements, Fábio Porfírio and Silvia Ladner for their help with the user study, the Future Learning Initiative of ETH Zurich for funding and support, the participants of the user studies for making our project possible, Dr. Felix Friedrich for hosting our study in his class, and the reviewers for their insightful comments.

REFERENCES

[1] Dor Abrahamson and Robb Lindgren. 2014. *Embodiment and embodied design.*

[2] Dor Abrahamson, Mitchell J Nathan, Caro Williams-Pierce, Candace Walkington, Erin R Ottmar, Hortensia Soto, and Martha W Alibali. 2020. The future of embodied design for mathematics teaching and learning. In *Frontiers in Education*, Vol. 5. Frontiers Media SA, 147.

[3] Michèle Artigue. 2009. The teaching and learning of mathematics at the university level. *Colección Digital Eudoxus* 7 (2009).

[4] Karla Atkins, Jiangzhuo Chen, VS Anil Kumar, and Achla Marathe. 2009. The structure of electrical networks: a graph theory based analysis. *International journal of critical infrastructures* 5, 3 (2009), 265–284.

[5] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[6] Daniel M Belenky and Lennart Schalk. 2014. The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review* 26, 1 (2014), 27–50.

[7] Trevor S Breusch and Adrian R Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the econometric society* (1979), 1287–1294.

[8] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[9] Jerome Seymour Bruner et al. 1966. *Toward a theory of instruction*. Vol. 59. Harvard University Press.

[10] Kira J Carbonneau, Scott C Marley, and James P Selig. 2013. A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology* 105, 2 (2013), 380.

[11] David J Chalmers. 2011. Verbal disputes. *Philosophical Review* 120, 4 (2011), 515–566.

[12] John M Chambers, Anne E Freeny, and Richard M Heiberger. 2017. Analysis of variance; designed experiments. In *Statistical models in S*. Routledge, 145–193.

[13] Julia Chatain, Virginia Ramp, Venera Gashaj, Violaine Fayolle, Manu Kapur, Bob Sumner, and Stéphane Magnenat. 2022. Grasping Derivatives: Teaching Mathematics through Embodied Interactions using Tablets and Virtual Reality. In *Interaction Design and Children (IDC'22)*.

[14] Julia Chatain, Danielle M Sisserman, Lea Reichardt, Violaine Fayolle, Manu Kapur, Robert W Sumner, Fabio Zünd, and Amit H Bermano. 2020. DigiGlo: Exploring the Palm as an Input and Display Mechanism through Digital Gloves. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 374–385.

[15] Wai-Kai Chen. 1997. *Graph theory and its engineering applications*. Vol. 5. World Scientific.

[16] R Breckinridge Church and Susan Goldin-Meadow. 1986. The mismatch between gesture and speech as an index of transitional knowledge. *Cognition* 23, 1 (1986), 43–71.

[17] M De Vega, A Graesser, and AM Glenberg. 2008. Framing the debate. *Symbols, embodiment, and meaning* (2008), 85–116.

[18] John Dewey. 1910. How We Think, DC Heath & Co. *Boston, Mass* 224 (1910).

[19] Zoltan Dienes. 2014. Using Bayes to get the most out of non-significant results. *Frontiers in psychology* 5 (2014), 781.

[20] Katinka Dijkstra and Lysanne Post. 2015. Mechanisms of embodiment. *Frontiers in Psychology* 6 (2015), 1525.

[21] Paul Dourish. 2004. *Where the action is: the foundations of embodied interaction*. MIT press.

[22] Adam Drogemuller, Andrew Cunningham, James Walsh, Maxime Cordeil, William Ross, and Bruce Thomas. 2018. Evaluating navigation techniques for 3d graph visualizations in virtual reality. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*. IEEE, 1–10.

[23] Ajendra Dwivedi, Xinghuo Yu, and Peter Sokolowski. 2010. Analyzing power network vulnerability with maximum flow based centrality approach. In *2010 8th IEEE International Conference on Industrial Informatics*. IEEE, 336–341.

[24] Thomas J Faulkenberry, Alexander Ly, and Eric-Jan Wagenmakers. 2020. Bayesian inference in numerical cognition: A tutorial using JASP. *Journal of Numerical Cognition* 6, 2 (2020), 231–259.

[25] Lester Randolph Ford Jr. and Delbert Ray Fulkerson. 1956. Maximal flow through a network. *Canadian Journal of Mathematics* (1956), 399–404.

[26] David Furió, Stéphanie Fleck, Bruno Bousquet, Jean-Paul Guillet, Lionel Canioni, and Martin Hachet. 2017. Hobit: Hybrid optical bench for innovative teaching. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 949–959.

[27] Emily R Fyfe, Nicole M McNeil, and Stephanie Borjas. 2015. Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction* 35 (2015), 104–120.

[28] Emily R Fyfe, Nicole M McNeil, Ji Y Son, and Robert L Goldstone. 2014. Concreteness fading in mathematics and science instruction: A systematic review. *Educational psychology review* 26, 1 (2014), 9–25.

[29] Emily R Fyfe, Bethany Rittle-Johnson, and Marci S DeCaro. 2012. The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology* 104, 4 (2012), 1094.

[30] Dedre Gentner and Donald R Gentner. 1983. Flowing waters or teeming crowds: Mental models of electricity. *Mental models* 99 (1983), 129.

[31] Renaud Gervais, Joan Sol Roo, and Martin Hachet. 2016. Tangible viewports: Getting out of flatland in desktop environments. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. 176–184.

[32] James J Gibson. 2014. *The ecological approach to visual perception: classic edition*. Psychology Press.

[33] Arthur Glenberg, Manuel De Vega, and Arthur C Graesser. 2012. Framing the debate. In *Symbols and embodiment: Debates on meaning and cognition*. Oxford University Press.

[34] Stevan Harnad. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42, 1-3 (1990), 335–346.

[35] Kristina Höök, Martin P Jonsson, Anna Ståhl, and Johanna Mercurio. 2016. Somaesthetic appreciation design. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3131–3142.

[36] Derek R Hopko, Rajan Mahadevan, Robert L Bare, and Melissa K Hunt. 2003. The abbreviated math anxiety scale (AMAS) construction, validity, and reliability. *Assessment* 10, 2 (2003), 178–182.

[37] Eva Hornecker. 2016. The to-and-fro of sense making: Supporting users' active indexing in museums. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 2 (2016), 1–48.

[38] Wen Huang. 2020. *Investigating the novelty effect in virtual reality on stem learning*. Ph. D. Dissertation. Arizona State University.

[39] Yi-Jheng Huang, Takanori Fujiwara, Yun-Xuan Lin, Wen-Chieh Lin, and Kwan-Liu Ma. 2017. A gesture system for graph visualization in virtual reality environments. In *2017 ieee pacific visualization symposium (pacificvis)*. IEEE, 41–45.

[40] Harold Jeffreys. 1961. *The theory of probability, 3rd edition*. Oxford, United Kingdom: Oxford University Press.

[41] Mina C Johnson-Glenberg and Colleen Megowan-Romanowicz. 2017. Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive research: principles and implications* 2, 1 (2017), 1–28.

[42] Jennifer A Kaminski, Vladimir M Sloutsky, and Andrew F Heckler. 2008. The advantage of abstract examples in learning math. *Science* 320, 5875 (2008), 454–455.

[43] Madeleine Keehner and Martin H Fischer. 2012. Unusual bodies, uncommon behaviors: individual and group differences in embodied cognition in spatial tasks. *Spatial Cognition & Computation* 12, 2-3 (2012), 71–82.

[44] John M Keller. 2010. The Arcs model of motivational design. In *Motivational design for learning and performance*. Springer, 43–74.

[45] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.

[46] Anita Körner, Sascha Topolinski, and Fritz Strack. 2015. Routes to embodiment. *Frontiers in psychology* 6 (2015), 940.

[47] Derek A Kuipers, Gijs Terlouw, Bard O Wartena, Job TB van't Veer, Jelle T Prins, and Jean Pierre EN Pierie. 2017. The role of transfer in designing games and simulations for health: systematic review. *JMIR Serious Games* 5, 4 (2017), e7880.

[48] Learning Lab. 2010. The Instructional Materials Motivation Survey (IMMS). https://learninglab.uni-due.de/research-instrument/13887. [Online; accessed 03-May-2022].

[49] George Lakoff and Mark Johnson. 2008. *Metaphors we live by*. University of Chicago press.

[50] George Lakoff and Rafael Núñez. 2000. *Where mathematics comes from*. Vol. 6. New York: Basic Books.

[51] Katharina Loibl, Ido Roll, and Nikol Rummel. 2017. Towards a theory of when and how problem solving followed by instruction supports learning. *Educational psychology review* 29, 4 (2017), 693–715.

[52] Nicole Loorbach, Oscar Peters, Joyce Karreman, and Michaël Steehouder. 2015. Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology. *British journal of educational technology* 46, 1 (2015), 204–218.

[53] Nicole M McNeil and Emily R Fyfe. 2012. "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction* 22, 6 (2012), 440–448.

[54] Edward F Melcer and Katherine Isbister. 2016. Bridging the physical divide: a design framework for embodied learning games and simulations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2225–2233.

[55] Edna Holland Mory. 2013. Feedback research revisited. In *Handbook of research on educational communications and technology*. Routledge, 738–776.

[56] Florian'Floyd' Mueller, Richard Byrne, Josh Andres, and Rakesh Patibanda. 2018. Experiencing the body as play. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[57] Charlotte H Müller. 2022. Facilitating Learning of Quantum Chemical Concepts through Grounding in Sensory Experience. In *General Proceedings of the ISLS Annual Meeting 2022*.

[58] Mitchell J Nathan. 2021. *Foundations of embodied learning: A paradigm for education*. Routledge.

[59] Mitchell J Nathan and Candace Walkington. 2017. Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive research: principles and implications* 2, 1 (2017), 1–20.

[60] Mitchell J Nathan, Candace Walkington, and Michael Swart. 2022. Designs for Grounded and Embodied Mathematical Learning. (2022).

[61] Don Norman. 2013. *The design of everyday things: Revised and expanded edition.* Basic books.

[62] Erin R Ottmar, Candace Walkington, Dor Abrahamson, Mitchell J Nathan, Avery Harrison, and Carmen Smith. 2019. Embodied Mathematical Imagination and Cognition (EMIC) Working Group. *North American Chapter of the International Group for the Psychology of Mathematics Education* (2019).

[63] Seymour Papert. 1980. The gears of my childhood. *Mindstorms: Children, Computers, and Powerful Ideas. Basic Books. USA* (1980).

[64] Siyou Pei, Alexander Chen, Jaewook Lee, and Yang Zhang. 2022. Hand Interfaces: Using Hands to Imitate Objects in AR/VR for Expressive Interactions. In *CHI Conference on Human Factors in Computing Systems.* 1–16.

[65] David N Perkins, Gavriel Salomon, et al. 1992. Transfer of learning. *International encyclopedia of education* 2 (1992), 6452–6457.

[66] Wim TJL Pouw, Tamara Van Gog, and Fred Paas. 2014. An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review* 26, 1 (2014), 51–72.

[67] Stacy J Priniski, Cameron A Hecht, and Judith M Harackiewicz. 2018. Making learning personally meaningful: A new framework for relevance research. *The Journal of Experimental Education* 86, 1 (2018), 11–29.

[68] Martina A Rau. 2017. Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review* 29, 4 (2017), 717–761.

[69] Wolff-Michael Roth. 2001. Gestures: Their role in teaching and learning. *Review of educational research* 71, 3 (2001), 365–392.

[70] Wolff-Michael Roth and Alfredo Jornet. 2013. Situated cognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 4, 5 (2013), 463–478.

[71] J Patrick Royston. 1982. An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 31, 2 (1982), 115–124.

[72] Michael Sambol. 2015. Ford-fulkerson in 5 minutes — step by step example. https://www.youtube.com/watch?v=Tl90tNtKvxs. [Online; accessed 03-May-2022].

[73] Anthony J Sanford. 2008. Defining embodiment in understanding. (2008).

[74] Luigi Sansonetti, Julia Chatain, Pedro Caldeira, Violaine Fayolle, Manu Kapur, and Robert W Sumner. 2021. Mathematics Input for Educational Applications in Virtual Reality. In *ICAT-EGVE 2021-International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments.* Eurographics Association.

[75] Alan H Schoenfeld. 2016. Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics (Reprint). *Journal of Education* 196, 2 (2016), 1–38.

[76] Rhea Sharma, Edward F Melcer, and Dominic Kao. 2022. *Exploring Relevance, Meaningfulness, and Perceived Learning in Entertainment Games.* Technical Report. EasyChair.

[77] Stephanie A Shields, Mary E Mallory, and Angela Simon. 1989. The body awareness questionnaire: reliability and validity. *Journal of personality Assessment* 53, 4 (1989), 802–815.

[78] Tanmay Sinha and Manu Kapur. 2021. When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research* 91, 5 (2021), 761–798.

[79] Katta Spiel. 2021. The bodies of tei–investigating norms and assumptions in the design of embodied interaction. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction.* 1–19.

[80] Kay M Stanney, Robert S Kennedy, and Julie M Drexler. 1997. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, Vol. 41. SAGE Publications Sage CA: Los Angeles, CA, 1138–1142.

[81] Heather A Stoner, Donald L Fisher, and Michael Mollenhauer. 2011. Simulator and scenario factors influencing simulator sickness. *Handbook of driving simulation for engineering, medicine, and psychology* (2011).

[82] Sangho Suh, Martinet Lee, and Edith Law. 2020. How do we design for concreteness fading? survey, general framework, and design dimensions. In *Proceedings of the Interaction Design and Children Conference.* 581–588.

[83] John Sweller. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction* 4, 4 (1994), 295–312.

[84] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. The split-attention effect. In *Cognitive load theory.* Springer, 111–128.

[85] Apostolos Syropoulos and Eleni Tatsiou. 2021. Vague mathematics. *Vagueness in the Exact Sciences: Impacts in Mathematics, Physics, Chemistry, Biology, Medicine, Engineering and Computing* (2021), 19–32.

[86] Sofia Tancredi, Julia Wang, Helen Tong Li, Carissa Jiayuan Yao, Genna Macfarlan, and Kimiko Ryokai. 2022. Balance Board Math:"Being the graph" through the sense of balance for embodied self-regulation and learning. In *Interaction Design and Children.* 137–149.

[87] Cathy Tran, Brandon Smith, and Martin Buschkuehl. 2017. Support of mathematical thinking through embodied cognition: Nondigital and digital approaches. *Cognitive Research: Principles and Implications* 2, 1 (2017), 1–18.

[88] Dragan Trninic, Manu Kapur, and Tanmay Sinha. 2019. The Disappearing "Advantages of Abstract Examples in Learning Math". In *Annual Conference of the Cognitive Science Society.*

[89] Johnny van Doorn, Don van den Bergh, Udo Böhm, Fabian Dablander, Koen Derks, Tim Draws, Alexander Etz, Nathan J Evans, Quentin F Gronau, Julia M Haaf, et al. 2021. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review* 28, 3 (2021), 813–826.

[90] Roy Wagner. 2019. Mathematical Abstraction as Unstable Translation Between Concrete Presentations. *Philosophy of Mathematics Education Journal* 35 (2019).

[91] Alan Weir. 2011. Formalism in the Philosophy of Mathematics. (2011).

[92] Rand R Wilcox. 2011. *Introduction to robust estimation and hypothesis testing.* Academic press.

[93] Uri Wilensky. 1991. *Abstract meditations on the concrete and concrete implications for mathematics education.* Epistemology and Learning Group, MIT Media Laboratory Cambridge, MA.

[94] Benedikt Wisniewski, Klaus Zierer, and John Hattie. 2020. The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology* 10 (2020), 3087.