Analyzing Self-Explanations in Mathematics: Gestures and Written Notes Do Matter

Alexander Salle*

Institute for Mathematics, School of Mathematics/Computer Science, Osnabrück University, Osnabrück, Germany

When learners self-explain, they try to make sense of new information. Although research has shown that bodily actions and written notes are an important part of learning, previous analyses of self-explanations rarely take into account written and non-verbal data produced spontaneously. In this paper, the extent to which interpretations of self-explanations are influenced by the systematic consideration of such data is investigated. The video recordings of 33 undergraduate students, who learned with worked-out examples dealing with complex numbers, were categorized successively including three different data bases: (a) verbal data, (b) verbal and written data, and (c) verbal, written and non-verbal data. Results reveal that including written data (notes) and non-verbal data (gestures and actions) leads to a more accurate analysis of self-explanations than an analysis solely based on verbal data. This influence is even stronger for the categorization of self-explanations as adequate or inadequate.

Keywords: self-explanation, gesture, multimodality, trigonometry, complex numbers

INTRODUCTION

Imagine a learner considering a worked-out example that presents the solution of a task dealing with right-angled triangles. The given example includes the complete solution of that task without describing the theorems or principles that were used for the calculations depicted. While reading the worked-out example, the learner thinks aloud: "*Ok, this triangle ABC,... the missing side... was calculated. And they made it... the Pythagorean Theorem, I think. This side here... then would be... uhm, square root of. yes, that works.*" When confronted with worked-out examples, texts or other instructional material, learners can learn in different ways with worked-out examples, e.g. superficially or thoroughly. The statement above illustrates an advantageous learning approach: The learner explains to herself the calculations in the material drawing on her (activated) prior knowledge of the Pythagorean Theorem, she is *self-explaining*.

Self-explaining means explaining something to oneself by generating information not provided in the given material and by creating inferences to organize given or new information in order to make sense of the material – it is a generative activity that occurs during learning (Chi, 2000, 2009; Fiorella and Mayer, 2016). Typical examples of self-explaining are activating and integrating prior knowledge, integrating different representations in a text, and clarifying requirements for

depicted mathematical operations (Chi et al., 1989; Chi, 2000). To distinguish cognitive processes and overt activities of learners, utterances generated during self-explaining—like the statement depicted above—are called *self-explanations* (Chi, 2000, 2009).

Self-explaining as a learning strategy can serve to aid learners' comprehension of a topic (de Koning et al., 2011; Chiu and Chi, 2014; Wylie and Chi, 2014; Fiorella and Mayer, 2016). There are several broadly approved methods to foster self-explaining like different forms of self-explanation prompts (e.g. Chi et al., 1994; Berthold et al., 2009; van der Meij and de Jong, 2011; Hefter et al., 2015) or trainings (McNamara, 2004; Kurby et al., 2012; Hodds et al., 2014), and it has been shown that these methods increase learning outcomes and understanding. In summary, when learning with instructional material, self-explaining is essential to a deep and meaningful understanding (Renkl, 2014; Wylie and Chi, 2014). Self-explanations, quantified and counted based on transcripts or recordings, have been identified as main predictors for learning outcomes in psychometrical designs when investigating the benefits of instructional material, e.g. worked-out examples or instructional texts (e.g. Chi et al., 1989; Butcher, 2006; Griffin et al., 2008).

These findings have been replicated in different domains. Especially, a plethora of studies focusing on instructional material and self-explanations has been conducted with mathematical or mathematically-related content (e.g. Chi et al., 1989; Renkl, 1997; Neuman and Schwarz, 1998; Stark, 1999; Renkl et al., 2004).

Until the 1980s, many researchers thought of mathematics as a purely cognitive and disembodied discipline, "based on the premise of a human mind-body split and of the transcendence of mind over body" (Gerofsky, 2014). Since then, many studies have shown that the whole body and its different modalities are important parts of the communication about and the learning of mathematics (Núñez et al., 1999; Arzarello et al., 2008; Radford, 2009; Gerofsky, 2010).

While there are numerous results of qualitative as well as quantitative studies underlining this importance, it remains unclear so far, how significant the inclusion of further modalities for quantifications of learning processes really is. In how far do individual differences, e.g. in gesturing behavior, influence the quantification of learning processes? Apart from the methodological level, these questions are also relevant when estimating privacy issues (audio recordings vs. video recordings) and necessary resources (duration of data analysis, expensive digital tools for data analysis) for empirical studies. Since quantifying self-explanations is an important and frequently used approach in studies analyzing meaningful learning, self-explaining is a fruitful concept for further investigations. Previous research on self-explanations mostly rely solely on verbal transcripts of think-aloud sessions (Ericsson and Simon, 1993; Chi, 1997) and exclude non-verbal or written data. In most cases, thinking aloud in self-explanation studies refers to the recording of verbal utterances, (i) without giving learners the opportunity to take notes, respectively, without considering written notes or their use during the learning sessions, and (ii) without considering gestures, actions or other non-verbal data during analysis.

The main question this paper intends to answer can, therefore, be stated as follows: *In what way does considering written notes and gestures influence the analysis of self-explaining in mathematical learning processes?*

Pursuing this question, an empirical study was conducted in which undergraduate students performed a learning exercise with worked-out examples from the field of complex numbers. Their learning processes were videotaped and categorized based on different data sources in order to analyze the influence of non-verbal and written data on the reconstruction of learning processes and self-explanations. Besides the results for self-explanations in particular, the design and the data analysis in this paper may be a transferable example for investigations that perform quantitative analyses to show the importance of gestures and notes in learning processes.

LITERATURE REVIEW

Self-Explanations

Adequate use of cognitive strategies and engaging in domain specific cognitive activities are essential to meaningful and deep learning (Swing and Peterson, 1988; Murayama et al., 2012). Since the seminal work of Chi et al. (1989), a large number of empirical works has shown the analytic and predictive power of self-explanations, which constitute a class of meaningful cognitive activities. Self-explanations are the main predictors for learning outcomes when investigating the benefits of instructional material like worked-out examples or instructional texts, and facilitate evaluation of instructional methods (e.g. Chi et al., 1989; Renkl, 1997, 2005; Butcher, 2006; de Koning et al., 2011).

Although both self-explaining and explaining are constructive activities (Chi, 2009), there are clear differences. While self-explaining is a cognitive activity that does not require verbalization (although it can be traced through overt activities and verbalization), explaining is inherently bound to communication. Self-explaining is based on one's own prior knowledge, whereas explaining must be based on the knowledge of another person. From a cognitive point of view, explaining needs additional selecting and organizing processes to give others a suitable and comprehensible explanation (Fiorella and Mayer, 2016). Typical examples of self-explanations are inferencing from depicted data (Wong et al., 2002), repairing misconceptions (Chi, 2000), explaining solution steps with prior knowledge (Chi et al., 1989), explaining the goals of an operation (Chi et al., 1989; Renkl, 1997; Renkl et al., 1998), and integrating symbolic calculations and iconic representations (Aleven and Koedinger, 2002). By definition, self-explanations can be incomplete, fragmented or even wrong, e.g. explaining the goals of an operation could address the wrong goals (Chi, 2000).

Self-explanation studies apply experiments in different domains, covering topics like the blood flow and the circulatory system (e.g. Ainsworth and Loizou, 2003; de Koning et al., 2011) or LISP programming (Pirolli and Recker, 1994; Bielaczyc et al., 1995). However, one of the domains most often investigated is mathematics. Such studies often deal with mathematical or mathematically-related content on a lower secondary level

like elementary probability theory (e.g. Renkl, 1997; Renkl et al., 2004), compound-interest calculations (e.g. Renkl et al., 1998), algebra word problems (e.g. Neuman and Schwarz, 1998; Neuman et al., 2000), elementary geometry (Wong et al., 2002), and Newton's laws and calculating forces (e.g. Chi et al., 1989; Chi and VanLehn, 1991). To determine the impact of self-explaining on test performance in those domains quantitatively, self-explanations are categorized and *quantified* based on data that is typically collected in one of two common procedures:[1]

(1) Coding of written texts produced by learners after or during their work with learning materials or tasks (e.g. Schworm and Renkl, 2006; Berthold et al., 2009);

(2) Coding of protocols from think-aloud sessions that are recorded during or after learners' work with instructional material or tasks (e.g. Chi et al., 1989; Renkl, 1997; Durkin and Rittle-Johnson, 2012; McEldoon et al., 2013).

In the first case, the participants themselves sum up their learning processes in writing. The resulting products are examined for passages which can be identified as self-explanations. In contrast to the second case, this procedure includes all available data (the written text).

During the second procedure, the participants, working individually, are instructed to think aloud while learning and working with different materials and/or tasks. Typically, the procedure follows the methodological principles of Ericsson and Simon (1993).

In the studies following the second procedure, participants in thinking aloud settings studies were audio- or videotaped. Subsequently, the recorded data was transformed into verbal protocols for further coding procedures.[2] When video data was collected, it was not considered in the analyses of verbal protocols (e.g. Bielaczyc et al., 1995; Neuman et al., 2000). If participants were allowed to take notes or sketch diagrams, these written documents were not analyzed synchronously with the verbal protocols and oftentimes not analyzed at all.[3]

Some authors try to gain insight into the role of incorrect self-explanations for learning (Wilkin, 1997; McNamara, 2004; Butcher, 2006; Ainsworth and Burcham, 2007; de Koning et al., 2011). Both negative and positive effects of incorrect self-explanations are revealed: For less demanding activities such as paraphrasing, the number of incorrect self-explanations correlates negatively with subsequently measured performance; for more demanding activities such as inferencing new information based on given texts, however, positive correlations

show the potential benefit of incorrect self-explanations (McNamara, 2004). These findings correspond with results that show the learning potential of incorrect self-explanations (Chi, 2000). However, the majority of studies on self-explanation do not distinguish between correct and incorrect self-explanations, although from a psychological as well as from a domain-specific perspective the differentiation would be an important issue for investigations of learning processes across all domains, especially for mathematics.

Learning and Multimodality

Many studies in the last 30 years have demonstrated that the whole body and its modalities are an important partaker of and a constitutive entity for communication and learning (Lakoff and Johnson, 1999; Núñez and Freeman, 1999; Gerofsky, 2014). Therefore, thoughts and language are created and expressed through many "modalities linked together – sight, hearing, touch, motor-actions, and so on." (Gallese and Lakoff, 2005, p. 456). Although especially mathematics is typically regarded as a highly cognitive discipline, many researchers have (re-)discovered and verified the body's importance for mathematical learning (Núñez et al., 1999; Lakoff and Núñez, 2000; Gerofsky, 2014). By emphasizing the importance of bodily modalities and their role for the origination of mathematics, gesturing can be seen as "a key element in communication and conceptualization" (Radford et al., 2009, p. 93).

As an important part of a multimodal perspective, analyses have shown teachers and learners gesturing frequently and intensely when communicating and thinking about mathematics (e.g. Alibali and DiRusso, 1999; Goldin-Meadow and Singer, 2003; Greiffenhagen and Sharrock, 2005; Edwards, 2008; Radford, 2009; Yoon et al., 2011; Kita et al., 2017). Gesturing can support uttered words as well as supplement or contradict them in different ways (e.g. Alibali and Goldin-Meadow, 1993; Alibali and DiRusso, 1999; Goldin-Meadow and Singer, 2003; Hegarty et al., 2005; Robutti, 2006; Arzarello et al., 2008). In combination with notes or other inscriptions, gestures are applied in specific and subtle ways to construct and communicate mathematical knowledge (Krause, 2016; Krause and Salle, 2016, and, more general, Streeck and Kallmeyer, 2001).

Some studies analyze the use of gestures during explanations or think-aloud settings. When learners explain things to each other or to a video camera, their expressions go beyond verbal utterances and are often accompanied by different kinds of gestures (Schwartz and Black, 1996; Emmorey and Casey, 2001; Hegarty et al., 2005; Salle, 2014). In about 50% of all think-aloud sessions with students who solved gear-problems, Schwartz and Black (1996) found that content-related "rotating" and "ticking" gestures could be observed. Hegarty et al. (2005) report that 98.5% of all identified verb phrases in a gear-problem experiment were accompanied by pointing and tracing gestures, revealing "important individual differences in the use of gesture in both communication and inference" (p. 354). Other works show how gestures are used in explanations to depict certain aspects of verbalized parts (Koschmann and LeBaron, 2002; Alibali and Nathan, 2012).

---

[1]Studies reviewed: Chi et al. (1989, 1994), Chi and VanLehn (1991), Pirolli and Recker (1994), Bielaczyc et al. (1995), Recker and Pirolli (1995), Renkl (1997, 2002), Neuman and Schwarz (1998), Renkl et al. (1998), Renkl et al. (2004), Stark (1999), Neuman et al. (2000), Wong et al. (2002), Ainsworth and Loizou (2003), McNamara (2004), Butcher (2006), Ainsworth and Burcham (2007), de Koning et al. (2011), Durkin and Rittle-Johnson (2012), Lin and Atkinson (2013), McEldoon et al. (2013), and Rittle-Johnson et al. (2015).

[2]An exception is Stark (1999) who applied a "live coding" during the intervention phase.

[3]Although some studies (e.g. Chi et al., 1989) allowed drawing and notetaking, no hints could be found that these documents were considered in the self-explanation data analysis.

Findings about the role of different modalities in self-explanation analyses are rare. In eye-tracking studies, learners' direction of gaze while integrating information given in material was analyzed; including data from eye-tracking devices allows for more accurate analyses of self-explanation (Merten, 2002; Conati and Merten, 2007; She and Chen, 2009; Hodds et al., 2014). However, no systematical analysis of self-explanations and the role of bodily modalities and inscriptions in think-aloud settings has been carried out yet. Thus, the extent to which the consideration of spontaneously produced written notes and non-verbal utterances, such as gestures, could influence the identification of self-explanations in think-aloud settings remains unclear. A multimodal analysis could help to identify adequate and inadequate self-explanations, improve explanations for learning gains or optimize the design of learning materials by providing better measures of self-explanation.

CONCEPTUAL FRAMEWORK AND RESEARCH QUESTIONS

Multimodality and Self-Explanations

An utterance is understood as an expressive product in the sense of the multimodal framework of Edwards and Robutti (2014). Hence, expressive products are "physical 'traces,' whether permanent or ephemeral, of people's actions" (p. 13). That includes speech and gestures as bodily based expressive products, and inscriptions like written words, symbols, graphs and visuals as external to the body (ibid.).

Applying this definition, spoken words as well as gestures and written products like sentences or drawings can "become fully partakers of the utterance itself" (Nemirovsky and Ferrara, 2008, p. 162). The different expressive products can be seen as facets of one single underlying mental process (Robutti, 2005; Edwards and Robutti, 2014) and, thus, allow identification of specific cognitive processes. Three main types of utterances (expressive products) can be distinguished:

- *Verbal utterances*: spoken sentences and words, shouts and other sounds.
- *Written utterances*: written inscriptions such as characters, words, sentences with specific syntax, drawings, figures, markers.
- *Non-verbal utterances*: gestures, sign language[4], facial expression, gaze, actions like the movements when writing or drawing.

Based on the remarks above, the definition of a self-explanation can be broadened. In the classical definition, self-explanations are defined as "units of utterances" produced by self-explaining (the cognitive activity), whereby utterances are meant to be verbal (Chi, 2000, p. 165). Hence from a multimodal perspective, a *self-explanation* will refer to *a unit of intertwined (verbal, written, and non-verbal) utterances produced by self-explaining*. This definition was used for the present study. All forms of verbal, written and non-verbal utterances

---

[4] A language 'spoken' through specific gestures (Sfard, 2008).

that were recorded on video are considered in this paper except for gaze and facial expression, which were not included in the analysis.

Gestures

Gestures are bodily based expressive products (Edwards and Robutti, 2014), cognitive processes are mirrored in speech and gesture. Gestures occur in combination with speech, but they also have self-oriented functions that may occur in combination with thought (Alibali et al., 2000; Kita et al., 2017). This paper follows McNeill (1992) and Kita et al. (2017) in their definition of gestures as idiosyncratic spontaneous movements of the hands and arms which depict action, motion, or shape, or indicate location or trajectory, they "include iconic gestures, metaphoric gestures, and deictic gestures" (Kita et al., 2017, p. 245) in the taxonomy described by McNeill (1992) and the differentiation of iconic gestures in mathematics as iconic-physical and iconic-symbolic formulated by Edwards (2008). As long as movement is not part of a functional act (taking notes, measuring something with a ruler), the movement is a gesture; otherwise, it is an action with a purpose (Goldin-Meadow, 2003; Kendon, 2004; Sabena, 2008).

In think-aloud scenarios, gestures can convey important information (Schwartz and Black, 1996; Hegarty et al., 2005; Yammiyavar et al., 2007); they are, therefore, co-thought and/or co-speech gestures. Following the definition of self-explanations from a multimodal perspective, a gesture as well as an action with or without simultaneous speech may allow a coder to identify self-explaining activities.

Inscriptions

An inscription is defined as "an external 'representation,' whether symbolic or imagistic, which is non-ephemeral and therefore amenable to reflection, review, and revision" (Edwards and Robutti, 2014). Since mathematics makes much use of external representations like symbols or graphs, inscriptions play an important role in doing, communicating and learning mathematics (Arzarello et al., 2011; Krause, 2016). Learners create and use inscriptions on paper or other mediums to store and highlight important information for themselves (Kiewra, 1989; Kobayashi, 2005, 2006), to organize them in specific ways (e.g. Eppler, 2006; Kenehan, 2007), or to use such collections when studying (Luo et al., 2016).

During think-aloud procedures, learners can refer to inscriptions already present in the instructional setting or produced by the learners themselves; therefore, researchers may identify self-explaining processes by considering inscriptions in combination with speech and gestures or without them.

Adequate and Inadequate Self-Explanations

Whether a self-explanation is "correct" or not depends not only on the utterance itself, but also on the content to be learned and the aims of the instructional setting. For example, the notion of a tangent line as a line that touches a circle at one point is absolutely adequate in elementary Euclidean geometry. When it comes to functions and calculus, however, this conception only holds true

for special cases and areas. Tangent lines on graphs of third grade polynomials might not fit this explanation.

Such examples illustrate that a classification into "right" or "wrong" self-explanations is difficult. Hence, a classification that distinguishes between *adequate* and *inadequate* self-explanations, always matched to the goals of an intervention and the instructional material itself, fits more precisely and is used throughout the paper.

Research Questions

Two research questions will guide the following analyses:

(1) Does the consideration of non-verbal utterances (e.g. spontaneous gestures and actions) and written utterances (e.g. notes and diagrams) alter or support the coding of self-explanations?
(2) Does the consideration of non-verbal and written utterances alter or support the determination of self-explanations as adequate or inadequate?

MATERIALS AND METHODS

Participants

The subjects were 33 undergraduate students at a German university (22 females, 11 males) who voluntarily participated in this study. The students ranged in age from 21 to 25 years ($M = 23.2$, $SD = 1.1$), all of them spoke German fluently. All participants were enrolled in teacher training courses for middle school mathematics at the time of the experiment. They were in the third, fourth or fifth semester of their course. All participants were familiar with worked-out examples and computers.

Materials

The participants worked with three worked-out examples that addressed the multiplication of complex numbers. The chosen topic was new and relevant to them: First, it concerned elementary concepts and objects like polynomials, the fundamental theorem of algebra and trigonometric functions. Dealing with complex numbers helps participants in understanding these contents, which will be relevant for teaching in school, from a more general point of view. Second, this topic represents a foundation for further lectures in algebra, analysis, geometry, etc. Third, the experience of becoming acquainted with a new number system has parallels to school children's first encounter with rational and real numbers, and, thus, gives future teachers the chance to reflect on certain aspects and obstacles concerning the encounter with new numbers.

Worked-out examples were chosen because they allow a structured investigation of self-explanations and constitute a common format in self-explanation research studies (see literature review). Every used worked-out example (**Figure 1**) was divided into three parts: (1) transformation of Cartesian coordinates into the trigonometric form of polar coordinates, (2) calculation of the product of two complex numbers

represented in the trigonometric form and (3) the geometrical representation of the calculated product.[5] The second worked-out example showed a second solution to a similar multiplication task and followed the same structure as the first example. In contrast to the first example, the coefficients of the complex numbers were fractions which represented vectors outside the first quadrant of the coordinate system. The third worked-out example dealt with trigonometrically represented polar coordinates that had to be transformed into Cartesian coordinates. Subsequent to this transformation, the Cartesian coordinates were multiplied and geometrically represented. Furthermore, the material contained different representations (symbolic calculations, geometrical representations and a part where those two were intertwined).

Objects of the Intervention

The following list provides a selection of favored self-explanations likely to arise during learning with the first worked-out example. They depict general principles of and insights into the mathematical topic and were results of a mathematical-content analysis[6] that was based on Hankel (1867) and Courant and Robbins (2010). The complete list forms the basis for the identification of adequate self-explanations (see **Supplementary Material** and section "Coding of Self-Explanations"). Numbers in parentheses refer to the three parts of the first example. A learner...

(1) explains the calculation of the vector's length as application of the Pythagorean Theorem.
(1) recognizes the calculation of the angle α as the application of a trigonometric equation in a right-angled triangle.
(1) integrates the symbolic representation of a complex number and its parts with respective characteristics of the geometrical counterpart.
(2) explains the change of the algebraic sign between line four and five of the calculation by the relation $i^2 = -1$.
(2) explains the simplification from line six to seven of the calculation as an application of the addition theorems.
(3) identifies the factor $\sqrt{74}$ as the length of the vector $s \cdot t$.
(3) recognizes that the resulting angle of $s \cdot t$ is the sum of α and β.

Procedure

First, participants had to complete a pre-test assessing their prior knowledge and competencies concerning complex numbers, trigonometric calculations and functions, and rules of calculating. The test contained 20 items where students had to draw complex numbers in different representations in coordinate systems (4), simplify simple and more complicated symbolic terms with roots and complex numbers (6), determine sine and cosine on right-angled triangles and in the unit circle (6) and give reasons for properties of real and complex numbers (4). The participants had 45 min to complete the test. On the one hand, the results revealed

---

[5]The worked-out examples were constructed based on state-of-the-art guidelines for instructional material (e.g. Mayer, 2014).
[6]This mathematical content analysis was carried out by a mathematician and two researchers of mathematics education, one of them being the author.

Task 1: Transform $s = 2 + 2i$ and $t = 0.5 + 3i$ into polar coordinates, multiply the complex numbers and plot s, t and $s \cdot t$ into a coordinate system.

Solution:

1. Transformation into polar coordinates: The length of s is calculated by
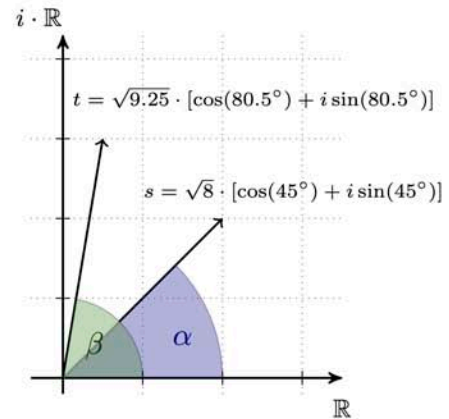
$$|s| = \sqrt{2^2 + 2^2} = \sqrt{8} \quad (\approx 2.83).$$

The angle between the $\mathbb{R}$-axis and the vector s is calculated by

$$\alpha = \tan^{-1}\left(\frac{2}{2}\right) = \tan^{-1}(1) = 45°.$$

Hence, the angle α amounts to $\alpha = 45°$, the magnitude of s is $|s| = \sqrt{8}$. Calculated in a similar way, the angle β of t is $\beta \approx 80.5°$ and the magnitude of t is $|t| = \sqrt{9.25}$. Hence, the polar coordinates are:

$$s = \sqrt{8} \cdot [\cos(45°) + i \cdot \sin(45°)] \quad \text{und} \quad t = \sqrt{9.25} \cdot [\cos(80.5°) + i \cdot \sin(80.5°)]$$

2. Multiply the complex numbers in polar coordinates:

$$\begin{aligned}
s \cdot t &= |s| \cdot (\cos\alpha + i \cdot \sin\alpha) \cdot |t| \cdot (\cos\beta + i \cdot \sin\beta) \\
&= \sqrt{8} \cdot [\cos(45°) + i \cdot \sin(45°)] \cdot \sqrt{9.25} \cdot [\cos(80.5°) + i \cdot \sin(80.5°)] \\
&= \sqrt{8} \cdot \sqrt{9.25} \cdot [\cos(45°) + i \cdot \sin(45°)] \cdot [\cos(80.5°) + i \cdot \sin(80.5°)] \\
&= \sqrt{8 \cdot 9.25} \cdot [\cos(45°)\cos(80.5°) + i \cdot \cos(45°)\sin(80.5°) + i \cdot \sin(45°)\cos(80.5°) + i^2 \cdot \sin(45°)\sin(80.5°)] \\
&= \sqrt{74} \cdot [\cos(45°)\cos(80.5°) + i \cdot \cos(45°)\sin(80.5°) + i \cdot \sin(45°)\cos(80.5°) - \sin(45°)\sin(80.5°)] \\
&= \sqrt{74} \cdot [\cos(45°)\cos(80.5°) - \sin(45°)\sin(80.5°) + i \cdot [\cos(45°)\sin(80.5°) + \sin(45°)\cos(80.5°)]] \\
&= \sqrt{74} \cdot [\cos(45° + 80.5°) + i \cdot \sin(45° + 80.5°)] \\
&(\approx 8.602 \cdot [\cos(125.5°) + i \cdot \sin(125.5°)])
\end{aligned}$$

3. Geometrical representation:

$$s \cdot t = \sqrt{74} \cdot [\cos(125.5°) + i \cdot \sin(125.5°)]$$
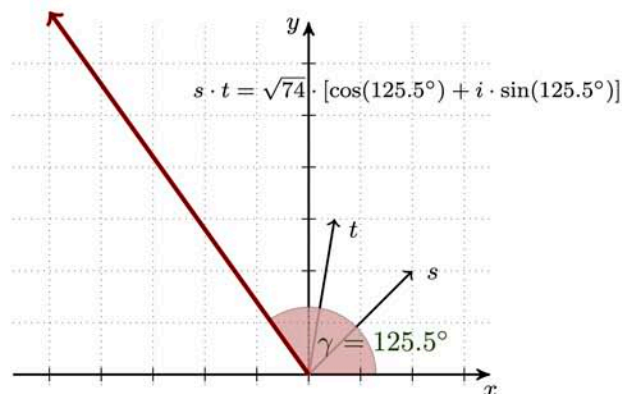
$$\gamma = 125.5°$$

FIGURE 1 | First worked-out example that was used in the intervention phase (translated from the German original, scaled-down version).

whether participants had the necessary basic arithmetic and algebraic knowledge for the intervention.[7] On the other hand, participants with too much experience in the field of complex numbers and polar coordinates could possibly be excluded from the study.[8]

In the intervention phase, the participants worked individually with three worked-out examples presented on paper. The assignment given to them was explained in the following way: "Try to understand the worked-out examples. Signal when you have finished. And please think aloud." After any period of 20 s of silence, the participants were reminded to think aloud. There was no time limit. The participants were permitted to use a prepared 'cheat sheet' with definitions and formulas, a pen, a triangle ruler and a calculator application on the computer screen. The think-aloud procedure followed the guidelines of Ericsson and Simon (1993) and Greene et al. (2011). Before the intervention phase there was a short training sequence for the think-aloud procedure. No guidelines on taking notes or gesturing were provided, so all occurrences of gestures and notes were produced spontaneously.

Data

The data base for the analysis of self-explanations consisted of video recordings from the intervention phases of 33 participants. These intervention sequences contained *verbal data* (participants' voices) and *non-verbal data* (recordings of the participants' upper bodies and bodily actions, recorded by the webcam on top of a computer display, and the recordings of calculations made on the computer screen). Furthermore, written notes, comments and calculations made on the worksheets, the 'cheat sheets' or additional blank paper were collected (referred to as *written data*).

Analysis

To analyze the data, a qualitative content analysis consisting of two phases was conducted (Mayring, 2000; Lamnek, 2010). The units of coding, context and analysis were refined successively.

Pilot Phase

During the pilot phase, preliminary category schemes were derived from a literature review (see section "Coding of Self-Explanations"). Based on two video sequences of students (unit of coding) who did not participate in the main study, the category schemes were revised, refined and adapted to the empirical findings (Mayring, 2000). Specifically, categories were extended to include written and non-verbal data, e.g. regarding the ways in which special examples of self-explanations were uttered through written or non-verbal expressive products, and regarding their intertwining with verbal utterances. The unit of context comprised a learning session of a participant, and the units of analysis comprised verbal utterances including gestures and notes. All codings were based on semantic features.

---

[7]A list summarizing necessary prior knowledge was another result of the mathematical content analysis mentioned above.
[8]One student was excluded.

Main Phase

Data gathered during the main phase consisted of video sequences from 33 participants. The units of context and analysis were identical to those employed during the pilot phase. In order to compare the self-explanations coded on the basis of different data and to follow common frameworks in self-explanation studies, recordings were divided into units of analysis based on verbal data. A unit of analysis was a sentence, a half-sentence or a shorter utterance, separated from other sentences by pauses. The segmentation of the data was done by both coders together. Disagreements were discussed and solved. This division was maintained throughout all coding procedures. The coding procedures are described in the next paragraph.

CODING OF SELF-EXPLANATIONS

The 33 video sequences were coded three times in consecutive coding procedures (described in sections "First Coding Procedure – Verbal Data," "Second Coding Procedure – Verbal and Written Data," and "Third Coding Procedure – Verbal, Written and Non-verbal Data"). Using the category schemes resulting from the pilot phase, each coding procedure was applied by a total of two coders familiar with the mathematical content and the research method. These coders coded the video material in all three procedures. 10% of the videos were encoded by both coders to determine the inter-coder reliability (described in section "Inter-Coder Reliability").The first procedure used only the verbal data from the video recordings to analyze verbal utterances, which included all spoken words, sentences and sounds. During the second procedure, the verbal data from the recordings and the written data were included. During the third procedure, the verbal, written and non-verbal data were included. Facial expressions were omitted from the analysis. The units of analysis were held constant throughout all procedures to allow for a one-to-one comparison between the analyses and thus facilitate quantitative comparisons. Examples of self-explanations identified during each of the three coding procedures can be found in the results section.

First Coding Procedure – Verbal Data

Based on the division into units of analysis, the verbal data was categorized first. The term "verbal data" refers to the audio track of the video recording, the first coding procedure was carried out based on this audio track. Two decisions were made for every unit of analysis: (a) Can a self-explanation be identified? (Identification of self-explanations). (b) If so, can the self-explanation be identified as adequate or inadequate with respect to the mathematical goals of the intervention? (Determination of adequate and inadequate self-explanations).

(a) *Identification of self-explanations*. Based on the verbal data, it was coded whether a unit of analysis was a self-explanation or not, using the following category scheme:

- (*Self-explanation*): A unit of analysis was coded as a self-explanation if a generation of inferences and/or a

mapping of inferences or information onto the learners' existing mental models could be identified; hence, if self-explaining could be reconstructed. Typical examples are activating prior knowledge for explanations of solution steps, calculations or representations; integrating different representations, e.g. symbolic and geometrical; and drawing inferences from information depicted in examples or the 'cheat sheet.'

- (*No self-explanation*): Examples for passages coded as no self-explanation are: reading a sentence from the instructional material without signs of bringing in new information through written or non-verbal utterances, or mentioning a number and pointing to it without integration of other information.

Category schemes used to develop the scheme for identification of self-explanations in the present study were those published by Chi et al. (1989), Renkl (1997), Wong et al. (2002), de Koning et al. (2011), and Salle (2014). The given categories were collected, merged if necessary and refined based on the results of the pilot phase, including written and non-verbal utterances. Finally, all identified categories were grouped together to the category "self-explanation" given above.

(b) *Identification of adequate and inadequate self-explanations.* Based on the verbal data, it was decided if a coded self-explanation was adequate or inadequate. If a self-explanation could not be identified as either, it was left unclassified in step (b). The following category scheme was used:

- (*Adequate*): Self-explanations that matched the goals of the intervention phase were coded as adequate self-explanations, e.g. if a learner identified the vector s as geometrical representation of the symbolically given complex number s (list "Objects of the intervention" in section "Materials and Methods").
- (*Inadequate*): With regard to the mathematical subject of complex numbers, self-explanations that revealed misconceptions, misunderstandings or inference errors were coded as inadequate, e.g. if a learner mistakenly explained the form of an equation by a multiplication with i instead of a rearrangement of terms.

(*Unclassified*): If it could not be derived whether the self-explanation was adequate or not, the self-explanation passage was labeled as unclassified.

Based on the mathematical content analysis described in Section "Materials and Methods," misconceptions and typical errors were determined. The formulation of the category scheme for adequate and inadequate self-explanations was based on Wilkin (1997), McNamara (2004), Butcher (2006), Ainsworth and Burcham (2007), and de Koning et al. (2011) regarding the remarks in Section "Adequate and Inadequate Self-Explanations." The categories of adequate and inadequate self-explanations were then supplemented with findings from the pilot phase.

Second Coding Procedure – Verbal and Written Data

The second coding procedure followed the same steps as the first, but was based on the verbal data from the recordings supplemented by the written data. To synchronize the two data sources, the written data was scanned and segmented by time before the second coding procedure. This preliminary segmenting was conducted by a person not involved in the coding procedures. For every unit of analysis, it was checked, based on the written and video data, which written expressive products were present at its end and its beginning. By this comparison, it could be derived which annotations were made during that unit. The set of these annotations constituted the written data belonging to the respective unit of analysis.[9]

The two decisions described above (identification of self-explanations and identification of adequate and inadequate self-explanations) are supported by a broader data base which includes the intertwining of the two types of utterances (see theoretical framework). Based on these two data sources, coding was carried out simultaneously without knowing the results of the first procedure. Differences between the results of the two coding procedures (units of analysis that were coded as self-explanations in one of the two procedures but not the other) were analyzed again by a second person to avoid coding errors. Discrepancies were then solved by consensus.

Third Coding Procedure – Verbal, Written and Non-verbal Data

For the third coding procedure, the data base was extended to verbal, written and non-verbal data. For every unit of analysis the coders listened to the audio data, read the segmented written products and, additionally, watched the video sequence of the participant's body, hands, and all objects on the table in front of the participants.

Although the coding procedure applied here is not based on *transcriptions* of verbal and non-verbal data, methodological advices from McNeill (2005, 263f.) were considered for the coding of the video sequences. Before beginning the coding procedures, the first pass to "facilitate interpretations of gesture productions on later passes," an analysis of the "product of elicitation" (results of the mathematical content analysis described in "Materials and Methods"), was discussed with all coders (McNeill, 2005, p. 264). Several purposes of the passes 5 and 6 described by Duncan and McNeill were considered for the coding procedure including non-verbal data to identify locations of gestures, beginning and ending of gesture phrases, and the identification of movements as gestures or actions. Hence, during the third coding procedure and based on the interplay of the three data sources, coders answered the following questions before they decided whether a unit of analysis could be counted as a self-explanation, respectively, as an adequate or inadequate self-explanation: is there a relevant movement visible in the non-verbal data? Is the relevant movement a gesture or an action? Is the whole gesture ("gesture phrase")

---

[9]This comparison was based on video data (see also the discussion section).

or action located inside the unit of analysis or does it begin/end in a adjacent unit? The respective answers provided the foundation for the categorization of passages as self-explanations and their classification as adequate or inadequate. Again, discrepancies in the three coding procedures as described above were analyzed by a different person to avoid errors and then solved by consensus.

Inter-Coder Reliability

Inter-coder reliability for all coding procedures was ascertained by two coders based on categorizations of 10% of the data. Because of the small number of categories, all reliability values were calculated with Cohen's κ, which takes random matches into account (Wirtz and Caspar, 2002). For the coding of self-explanations, the inter-coder reliability was 86.96% for the coding procedure based on verbal data, 88.93% for the procedure based on verbal and written data, and 89.2% for the coding based on all available data. Decisions on adequate and inadequate self-explanations showed an inter-coder reliability of 84.79% based on verbal data, 88.6% based on verbal and written data, and 90.11% for the coding based on all available data. Discrepancies were resolved by consensus.

RESULTS

The results section consists of four parts. Subsections "Coding of Self-Explanations," "Changes of the Coding Results With Different Data Bases," and "Alteration of the Interpretation of a Self-Explanation" depict results concerning the first research question (see section "Research Questions"): in what way will the consideration of non-verbal utterances (e.g. spontaneous gestures and actions) and written utterances (e.g. notes and diagrams) alter or support the coding of self-explanations? Subsection "Adequate and Inadequate Self-Explanations" depicts results concerning the second research question: In what way will the consideration of non-verbal and written utterances alter or support the determination of self-explanations as adequate or inadequate?

The pre-test results of the 33 subjects range between 36% and 87% (**Figure 2**). All participants had the basic knowledge that was necessary for the intervention phase (basic knowledge of sine and cosine, calculations with sine, cosine and real functions, and rearranging equations). There are small correlations between pre-test results and the number of coded self-explanations that increase with additional data sources (**Table 1**). Therefore, a strong relationship between prior knowledge and coding results could be ruled out.
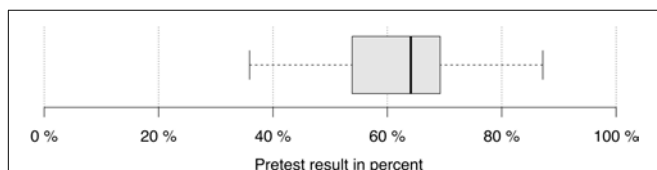

FIGURE 2 | Pre-test results of all participants.

TABLE 1 | Correlations (Bravais–Pearson) between pre-test results and coding results.

	Verbal	Verbal and written	Verbal, written, and non-verbal
Pretest result in %	0.17	0.21	0.27

Coding of Self-Explanations

On average, each one of the 33 data sessions lasted about 24:36 min (standard deviation: 14:28 min). In total, 935 self-explanations were coded based on all available data. Without non-verbal data and based on verbal and written data only, 738 self-explanations were coded, which amounts to a difference of 197 self-explanations. From the 197 self-explanations 31 could be coded because of actions like using the calculator or the 'cheat sheet.' The remaining 166 could be coded because of gestures and their intertwining with verbal and written data. In all but three of these 166 passages, participants used gestures that point to or retrace objects depicted in the material. Based on verbal data only, 676 self-explanations were coded; this amounts to a difference of 62 self-explanations that were not coded because written data were not considered. Compared to the coding results based on all available data, 259 self-explanations were not coded based on verbal data (**Figure 3**).

The individual results of the self-explanation coding show differences in the number of self-explanations coded during each of the three coding procedures (**Figure 4**). Some participants' self-explanations were coded based nearly exclusively on verbal data (no. 8, 16, 17, 26 in **Figure 4**), while others show more codings based on all available data (no. 1, 10, 11, 24, 25, 33 in **Figure 4**). Other learners showed greater numbers of self-explanations that were coded based on verbal and written data (no. 3, 10 in **Figure 4**).

The following sequence gives an example for a self-explanation that could only be coded during the coding procedure based on all available data.

A self-explanation only coded due to the inclusion of non-verbal data. Leo (participant no. 24 in **Figure 4**) is reading the lines of the calculation on the first example sheet (see **Figure 5**). The transcript depicts two units of analysis, divided by the pause in line 2.

Leo: At first cosine times cosine (points at #1 and #3) sine times co- (points at #2 and #3).,
ah, at first cosine times sine, (points at #1 and #4) sine times cosine (points at #2 and #3), and sine times sine (points at #2 and #4) with the i square.

Looking only at the verbal protocol without considering the accompanying pointing gestures, it seems as if Leo reads aloud the third line in the snippet, because the words refer almost exactly to the depicted formula (**Figure 5**). However, the accompanying pointing gestures intertwined with the verbal utterances reveal a self-explanation: She connects line two (the line she is pointing at) and line three (the line she compares her words to) of the calculation by carrying out the expansion of the product with her fingers. Hence, based on all available data,
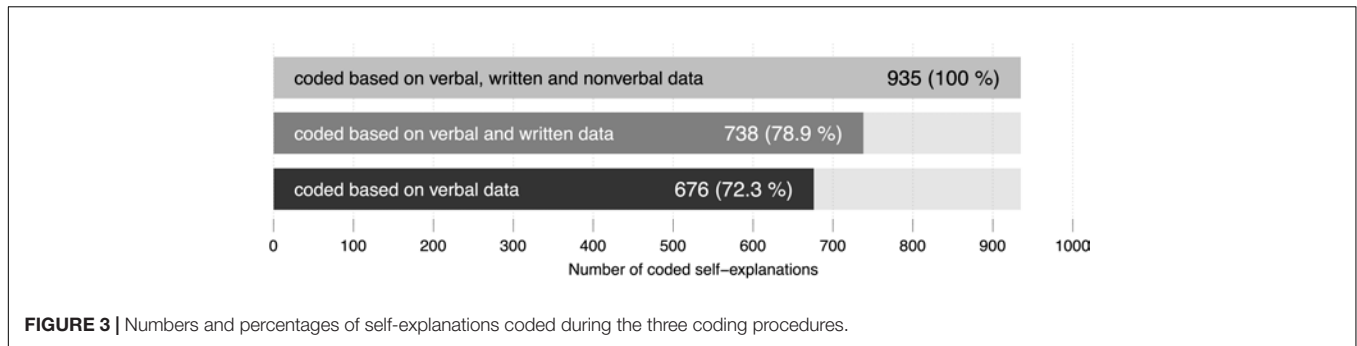
FIGURE 3 | Numbers and percentages of self-explanations coded during the three coding procedures.
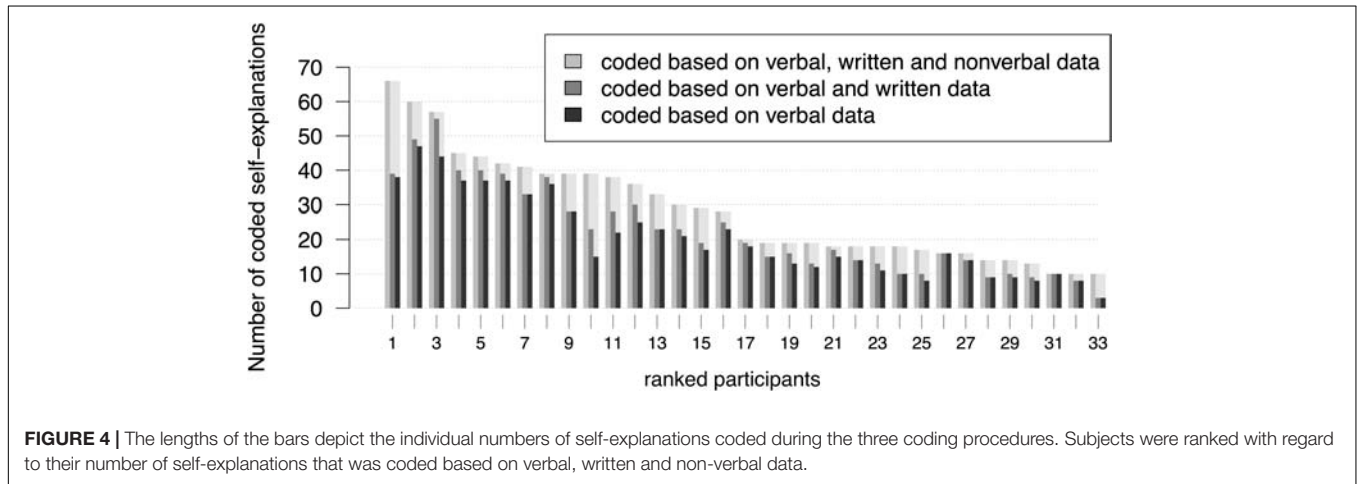

FIGURE 4 | The lengths of the bars depict the individual numbers of self-explanations coded during the three coding procedures. Subjects were ranked with regard to their number of self-explanations that was coded based on verbal, written and non-verbal data.
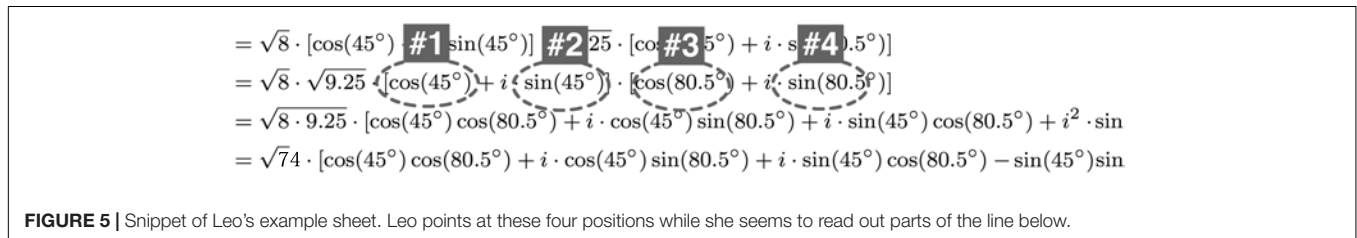

FIGURE 5 | Snippet of Leo's example sheet. Leo points at these four positions while she seems to read out parts of the line below.

this segment could be coded as a self-explanation in the third coding procedure. During this scene, Leo did not produce written data. This example illustrates a behavior frequently observed during the coding procedures. Verbal utterances that seem to be read aloud text passages become self-explanations when taking accompanying gestures into account.

Changes of the Coding Results With Different Data Bases

On average, 72.3% of all coded self-explanations were coded based on verbal data. Thus, there is an average difference of 28.14% in contrast to the coding based on all available data (see **Figure 3**). The individual differences between results based on these two data bases vary with a standard deviation of 16.19%. Individual results describing the proportion of self-explanations that were coded based on all available data range between 70% (learner no. 33 in **Figure 4**: seven out of ten self-explanations were coded based on all available data) and 0% (learners no. 26

and no. 31 in **Figure 4**: all self-explanations were coded based on verbal data only). Although there is a tendency toward higher differences in higher ranks, higher as well as lower differences can be found across the whole spectrum of ranks.

Although some partial sequences remain the same, there is a non-negligible difference between the different ranks and the different total numbers of self-explanations that could affect, e.g. subsequent statistical calculations predicting individual achievement (see **Figure 6**).

Alteration of the Interpretation of a Self-Explanation

The quantitative coding does not consider the concrete interpretation of a self-explanation, e.g. whether a coded self-explanation is an activation of prior knowledge or the identification of a goal of an operation. However, a detailed analysis of certain self-explanations reveals a further aspect regarding the influence of written and non-verbal data on
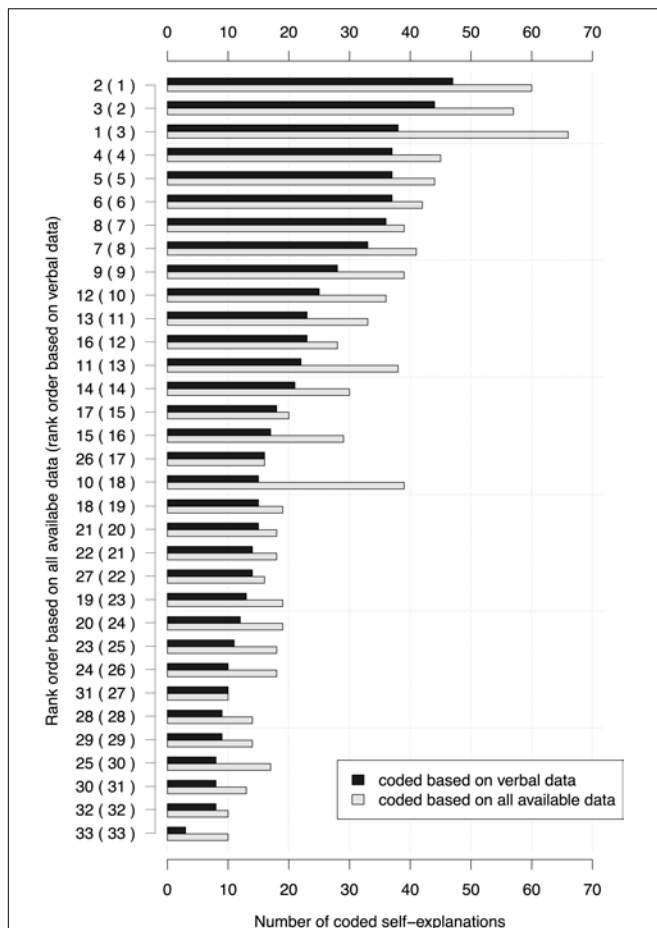
FIGURE 6 | Results of coding self-explanations with different data bases. The black bars depict the number of coded self-explanations based on verbal data. The light gray bars depict the number of coded-self-explanations based on verbal, written and non-verbal data. Bars are ordered according to the numbers of self-explanations coded based on verbal data (numbers in parentheses). In comparison, the order according to the numbers of self-explanations coded based on verbal, written and non-verbal data is depicted.

the identification of self-explanations. The following snippet shows in detail how a self-explanation receives an additional meaning when it is interpreted based on all available data. Lena (participant no. 10 in **Figure 4**) is working on the second example and tries to understand the transformation of $s = -\frac{1}{2} + \frac{1}{2} \cdot i$ into polar coordinates (see **Figure 7**).

Lena: Now I ask myself why, where this two by two comes from, (points at #1) but, .ah, exactly, this is here (points at #2) (points at #3). opposite side (lifts the pencil and points at #3 again). by adjacent side.

This self-explanation was coded based on verbal data because the words "opposite side" and "adjacent side" are neither depicted on the worked-out example nor on the 'cheat sheet.'[10]

[10]In the original protocol, she uses the words "Gegenkathete" for opposite side and "Ankathete" for adjacent side, which are technical terms in German.

So, without considering the pointing gestures, this self-explanation would be described as an activation of prior knowledge because Lena uses the technical terms she has learned in class. However, considering the pointing gestures leads to an additional interpretation as an integration of geometric and symbolic representations, because the participant combines the fraction and coordinate system using her words and gestures simultaneously. Generally, the additional data often revealed integrations of representations (as can be seen in Lena's protocol) or links between different information.

Results so far have focused on whether a unit of utterances is a self-explanation or not and how a self-explanation can be interpreted. In the following subsection, the influence of the different data on the categorization of self-explanations as adequate and inadequate will be analyzed.

Adequate and Inadequate Self-Explanations

Of all 935 self-explanations, 835 could be determined as adequate and 68 as inadequate. In the end, 32 self-explanations could not be identified as adequate or inadequate and were left unclassified.

Number of Coded Adequate and Inadequate Self-Explanations

835 adequate self-explanations could be identified by considering all available data. Based on verbal and written data, only 59.9% (500 of 835 self-explanations) were coded as adequate. About 48% (400 of 835 self-explanations) were coded as adequate based on verbal data (**Figures 8A,B**). Comparing the coding procedure based on verbal data and the coding procedure based on all available data, individual coding results of adequate self-explanations vary with a standard deviation of 20.55% around the mean of 48% (**Figure 9**). The individual proportion of adequate self-explanations coded based on all available data ranges between 84.8% (learner no. 10: only 5 of 33 adequate self-explanations were coded based on verbal data) and 0% (learners no. 26 and no. 31: all coded adequate self-explanations were coded based on verbal data). Thus again, although there is a tendency toward higher differences in higher ranks, higher differences as well as lower differences can be found across the whole spectrum of ranks. The overall percentages are even more distinct when reconstructing inadequate self-explanations: 68 inadequate self-explanations could be identified by considering all available data ($n = 68$). Based on verbal and written data, 47.1% (32) self-explanations were categorized as inadequate. Considering verbal data, only 26.5% (18) inadequate self-explanations could be identified.

To give an insight into the importance of non-verbal data for the coding of inadequate self-explanations, a typical example is presented in the following paragraph.

An Example of an Inadequate Self-Explanation Coded Based on Verbal and Non-verbal Data

Emily (participant no. 1 in **Figure 4**) reads out the fourth line of the solution: "The angle between the R-axis and the vector s is." (**Figure 7**).

1. Transformation into polar coordinates: The length of s is calculated by

$$|s| = \sqrt{2^2 + 2^2} = \sqrt{8} \quad (\approx 2.83).$$

The angle between the $\mathbb{R}$-axis and the vector s is calculated by

$$\alpha = \tan^{-1}\left(\frac{2}{2}\right) = \tan^{-1}(1) = 45°.$$

Hence, the angle of s amounts to $\alpha = 45°$, the magnitude of s is $|s| = \sqrt{8}$. Calculated in a similar way, the angle β of t is $\beta \approx 80.5°$ and the magnitude of t is $|t| = \sqrt{9.25}$. Hence, the

$t = \sqrt{9.25} \cdot [\cos(80.5°) + i\sin(80.5°)]$

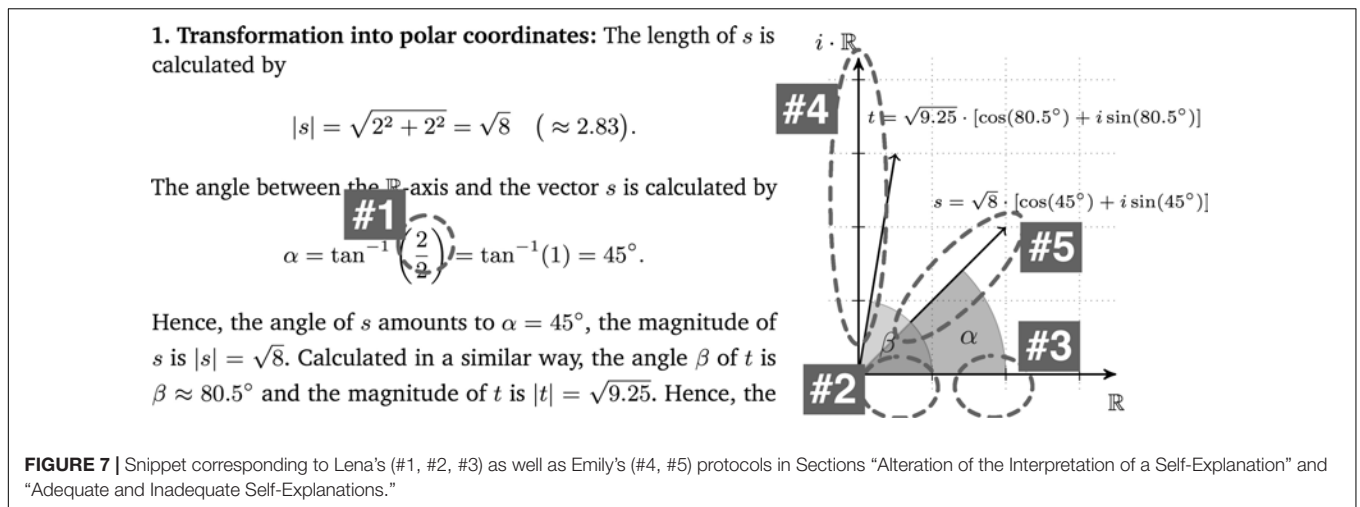$s = \sqrt{8} \cdot [\cos(45°) + i\sin(45°)]$

FIGURE 7 | Snippet corresponding to Lena's (#1, #2, #3) as well as Emily's (#4, #5) protocols in Sections "Alteration of the Interpretation of a Self-Explanation" and "Adequate and Inadequate Self-Explanations."
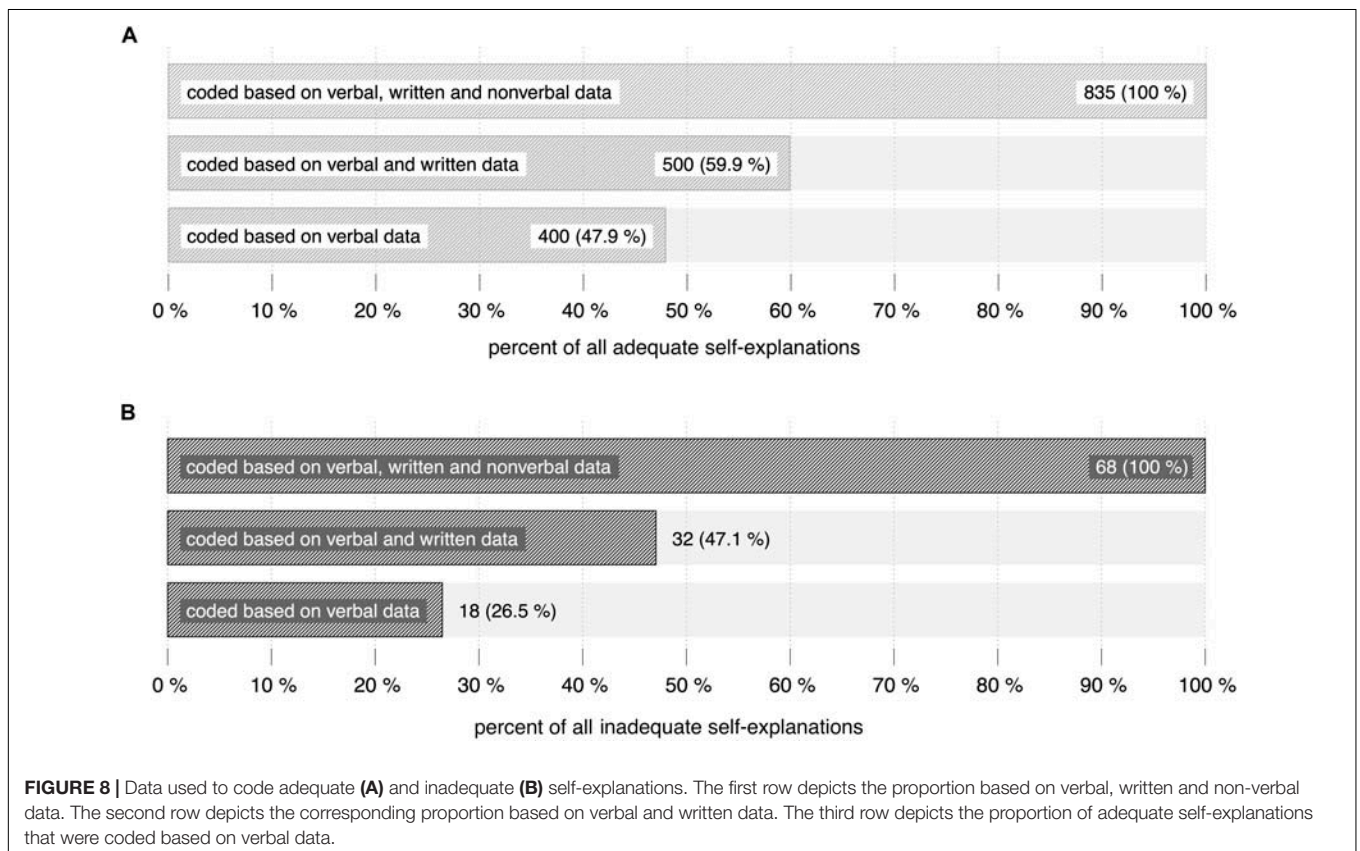

FIGURE 8 | Data used to code adequate **(A)** and inadequate **(B)** self-explanations. The first row depicts the proportion based on verbal, written and non-verbal data. The second row depicts the corresponding proportion based on verbal and written data. The third row depicts the proportion of adequate self-explanations that were coded based on verbal data.

Emily: Ok, the angle between the R-axis,. that's this one (traces along the i·R-axis at #4). and the vector s, this one. (traces along vector s at #5).

Without the pointing gestures, the coder does not know which part of the example her comment "that's this one" refers to. From the intertwining of speech and gesture it can be recognized that she is identifying the wrong axis. In general, self-explanations like linking of concepts or notions to geometrical representations are often acc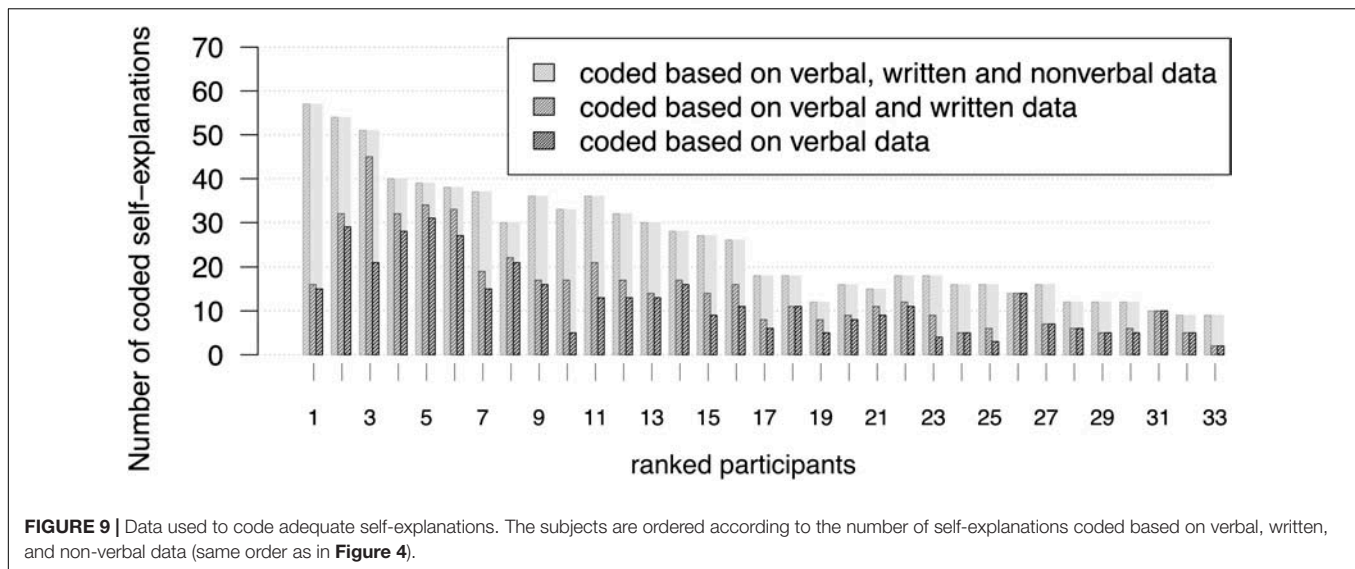ompanied by written notes or pointing gestures. In order to decide whether a self-explanation is adequate or not, coders have to know the positions, terms and numbers in the given calculation to which participants may refer.

DISCUSSION

Summary of the Findings

Although more than two-thirds of all self-explanations (676 of 935) could be reconstructed solely based on verbal

FIGURE 9 | Data used to code adequate self-explanations. The subjects are ordered according to the number of self-explanations coded based on verbal, written, and non-verbal data (same order as in **Figure 4**).

data, the consideration of gestures, actions and written products notably affected individual results distinctly: individual differences in self-explanations coded in the three procedures vary independently of the amount of self-explanations and the resulting ranks. The results underline the key role of video data to link the instructional material and written notes to verbal utterances: taking into account more data results in more distinct changes in the ranking. An analysis including written and non-verbal data seems to fit better to the concept of self-explaining than an analysis without.

The influence of written and non-verbal data is even more important for the coding of self-explanations as adequate or inadequate. Verbal data often leaves the researcher uncertain with regard to objects referred to and connections emphasized by participants. Taking into account the intertwining of verbal utterances, written notes and gestures seems to improve the reconstruction of adequate and inadequate self-explaining and, hence, complex learning processes.

The variability of individual coding results for self-explanations as well as adequate and inadequate self-explanations may have several reasons. Important factors why people gesture more or less are the spoken language (Pika et al., 2006), differences in cognitive abilities (e.g. Hostetter and Alibali, 2007; Wartenburger et al., 2010) or other personal characteristics (Hostetter and Potthoff, 2012). Apart from general differences in gesture frequency, processing difficulties and the strength of mental representations may have an influence on how frequently learners use gestures (Melinger and Kita, 2007; Sassenberg and van der Meer, 2010). Another explanation for the differing coding results based on written data might be the differing use of learning strategies that need paper and pencil, like note-taking (Kiewra et al., 1995; Kobayashi, 2006). Last but not least, the general use of self-explaining activities influence the individual coding results as well; empirical studies show that students have different

self-explanation-styles and, therefore, the quality and quantity of spontaneously generated self-explanations may differ distinctly (Renkl, 1997).

Methodological Discussion
The analyses described in this paper tried to thoroughly investigate multimodality and self-explanations, which is a very effortful endeavor. In order to facilitate similar methodological approaches, e.g. the application of tablets could decrease the effort: participants' written products could be segmented and assigned to specific units of analysis automatically. By omitting facial expressions, the current analysis of non-verbal data was less complex than an analysis including such expressions; however, considering such data could increase the precision of self-explanation analyses. The same holds true for gaze analysis via eye-tracking devices, which may give further insights into learners' self-explaining activities (e.g. Conati and Merten, 2007; Hodds et al., 2014).

This paper follows a traditional approach to self-explanation analysis as applied in many previous studies (e.g. Chi et al., 1989; Renkl, 1997; McEldoon et al., 2013) and extends this approach by including further modalities. Hence, on the one hand, the constant segmentation used throughout the study provides a reliable frame for a one-to-one comparison of analyses considering different data bases. Furthermore, the results offer some perspectives for methodological variations of traditional self-explanation studies. For example, by consideration of written and non-verbal data, maybe even sparse verbal protocols may reveal self-explaining activities (such sparse protocols were mentioned, e.g. in Renkl, 2002; Renkl et al., 2004). On the other hand, certain characteristics of multimodality could not be considered in the analysis, e.g. in what way different modalities alter the segmentation of the data or whether there are fundamental differences between self-explanations

accompanied by gestures or actions and self-explanations unaccompanied by gestures or actions. The use of gestures could be related to the proximity of uttered concepts to practical actions (Kita et al., 2017) or to the use of metaphors that express spatial concepts (Lakoff and Johnson, 1980; Lakoff and Núñez, 2000). Other explanations may refer to the cognitive effort – the more demanding cognitive processes are, e.g. conceptualization, the more often gestures occur (Hostetter et al., 2008) – or to the durability of self-explanations: frequent gesturing will "make learning last" (Wagner Cook et al., 2008).

The fact that in the present study 676 self-explanations could be coded solely based on verbal data does not imply the absence of gestures. Although not necessary for the coding decision, more than two thirds of these 676 self-explanations and about 80% of all 935 coded self-explanations were accompanied by gestures or actions.[11] Thinking aloud appears to be strongly connected to gesturing and acting in silence (Schwartz and Black, 1996; Hegarty et al., 2005), particularly if no visible human person is available as a dialogue partner (Emmorey and Casey, 2001).

More detailed analyses regarding the interplay of and emphasis on the modalities involved could provide deeper insights into the learners' cognitive effort of learners (e.g. more frequent use of gestures when learners engage in more demanding cognitive processes), their specific use of gestures in combination with speech and writing (e.g. lowering cognitive load by locating things with gesture or writing things down; intertwining use of writing and gesturing for themselves) and, more generally, into their generative cognitive activities. In combination with the broadened concept of self-explanation introduced in this paper, answers to these questions may provide a foundation for an intensified conceptual discussion and for the integration of self-explanations and gestures (Alibali and Goldin-Meadow, 1993; Kita et al., 2017). Comparisons to gesture-speech mismatches (Alibali and Goldin-Meadow, 1993), growth points (McNeill, 2002) or, more general, cognitive functions of gesturing like the activation, the manipulation, the packaging or the exploration of spatio-motoric information (Kita et al., 2017) offer promising starting points for further theoretical integration.

Implications for Learning and Instruction

Since instructional learning is an important method in schools and universities around the world, the implementation of self-explaining in such settings is highly relevant (Chiu and Chi, 2014). Considering the results presented in this paper, teachers seeking to identify specific self-explaining processes applied by learners have to carefully consider not only verbal utterances, but all multimodal aspects. Especially the subtle changes expressed by writing, gesturing and their combined use can help determine concrete self-explaining processes carried out more precisely (e.g. distinguishing reading of a passage, activating prior knowledge and integrating different representations) in order to facilitate individual learning processes. Based on the qualitative analyses in this

study, it could also be hypothesized that a classification of high quality and low quality self-explanations would be affected distinctly by the consideration of more modalities (cf. Roy and Chi, 2005). This could be of some benefit especially to the identification of successful self-explaining activities and elicitations through self-explanation prompts or trainings in the classroom, and could hereby confirm or alter the scope of such interventions to foster meaningful learning (Chi et al., 1994; Berthold et al., 2009).

Furthermore, teachers often support their explanations by writing, diagrams on a blackboard or a slideshow, and gestures (Alibali et al., 2014; Yeo et al., 2017). Since these teachers' gestures influence the learners' understanding of learners and are partly imitated by them (e.g. Goldin-Meadow et al., 1999; Cook and Goldin-Meadow, 2006), teachers' awareness of how such gestures influence learners' explanations and self-explaining could lead to a deeper understanding of what's going on in the learners' minds. Additionally, for learners as well as teachers, knowing in what way gesturing (to oneself and to others) can influence self-explaining would be of extremely high value for the understanding and improvement of learning. Deeper analyses of the functions and purposes of gestures during self-explaining could, for example, clarify to what extent actions and gestures may decrease cognitive load during learning (Goldin-Meadow, 2010; Alibali et al., 2011; Krause and Salle, 2016; Kita et al., 2017). This information could provide learners with the opportunity to optimize their self-explaining by an adequate use of specific gesturing.

Implications for Further Research

Several questions arise from the results presented. This paper does not address the influence of different coding results on predictions of achievement; to confirm the predictive power of multimodal self-explanations, further studies may quantify the extent to which those predictions vary by the inclusion of additional data sources during analysis. Therefore, multimodal self-explanations may serve as reliable measures for the construction and evaluation of instructional materials (e.g. Renkl, 2005; Butcher, 2006; de Koning et al., 2011). Furthermore, future analyses may reveal the extent to which gestures accompanying self-explanations can be characterized in terms of iconic, metaphorical and deictic gestures (McNeill, 1992; Goodwin, 2003; Edwards, 2008), or, in more detail, as tracing gestures or pointing gestures (cf. Hegarty et al., 2005). Additionally, it remains unclear to what extent these gestures can be found when dealing with geometrical representations instead of symbolic calculations, in what way they are intertwined with simultaneous verbal and written utterances and whether the presence or absence of the learning material and the presentation on paper or screen influence gesturing and learning. Regarding current theories on cognitive functions of gestures, answers to these questions could shed light on more fundamental questions regarding the use of gestures during thinking (Kita et al., 2017).

Although the results discussed here are specific to the mathematical domain, it could be hypothesized that written and

---

[11] Details of this analysis are not depicted in this paper.

non-verbal data may influence the results of self-explanation studies in other domains, too. The analytical approach therefore may provide a promising starting point for a deeper analysis of cognitive activities in general.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The patients/participants provided their written informed consent to participate in this study, which satisfied the ethical criteria and dispositions of our institution.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.513758/full#supplementary-material

Supplementary Data Sheet 1 | List of self-explanations.

Supplementary Data Sheet 2 | Worked-out examples.

REFERENCES

Ainsworth, S., and Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learn. Instr.* 17, 286–303. doi: 10.1016/j.learninstruc.2007.02.004

Ainsworth, S., and Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cogn. Sci.* 27, 669–681. doi: 10.1207/s15516709cog2704_5

Aleven, V., and Koedinger, K. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cogn. Sci.* 26, 147–179. doi: 10.1207/s15516709cog2602_1

Alibali, M. W., and DiRusso, A. A. (1999). The function of gesture in learning to count: more than keeping track. *Cogn. Dev.* 14, 37–56. doi: 10.1016/s0885-2014(99)80017-3

Alibali, M. W., and Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: what the hands reveal about a child's state of mind. *Cogn. Psychol.* 25, 468–523. doi: 10.1006/cogp.1993.1012

Alibali, M. W., Kita, S., and Young, A. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Lang. Cogn. Process.* 15, 593–613. doi: 10.1080/016909600750040571

Alibali, M. W., and Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: evidence from learners' and Teachers' Gestures. *J. Learn. Sci.* 21, 247–286. doi: 10.1080/10508406.2011.611446

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Breckinridge Church, R., Jacobs, S. A., Knuth, E. J., et al. (2014). How teachers link ideas in mathematics instruction using speech and gesture: a corpus analysis. *Cogn. Instr.* 32, 65–100. doi: 10.1080/07370008.2013.858161

Alibali, M. W., Spencer, R. C., Knox, L., and Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychol. Sci.* 22, 1138–1144. doi: 10.1177/0956797611417722

Arzarello, F., Ascari, M., and Sabena, C. (2011). A model for developing students' example space: the key role of the teacher. *ZDM Math. Educ.* 43, 295–306. doi: 10.1007/s11858-011-0312-y

Arzarello, F., Paola, D., Robutti, O., and Sabena, C. (2008). Gestures as semiotic resources in the mathematics classroom. *Educ. Stud. Math.* 70, 97–109. doi: 10.1007/s10649-008-9163-z

Berthold, K., Eysink, T. H. S., and Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instr. Sci.* 37, 345–363. doi: 10.1007/s11251-008-9051-z

Bielaczyc, K., Pirolli, P., and Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: investigating the effects of knowledge acquisition activities on problem solving. *Cogn. Instr.* 13, 221–252. doi: 10.1207/s1532690xci1302_3

Butcher, K. (2006). Learning from text with diagrams - promoting mental model development and inference generation. *J. Educ. Psychol.* 98, 182–197. doi: 10.1037/0022-0663.98.1.182

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *J. Learn. Sci.* 6, 271–315. doi: 10.1207/s15327809jls0603_1

Chi, M. T. H. (2000). "Self-explaining expository texts: the dual processes of generating inferences and repairing mental models," in *Advances in Instructional Psychology*, ed. R. Glaser (Mahwah, NJ: Lawrence Erlbaum Associates), 161–238.

Chi, M. T. H. (2009). Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Top. Cogn. Sci.* 1, 73–105. doi: 10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., and Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cogn. Sci.* 13, 145–182. doi: 10.1207/s15516709cog1302_1

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., and LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cogn. Sci.* 18, 439–477. doi: 10.1016/0364-0213(94)90016-7

Chi, M. T. H., and VanLehn, K. (1991). The content of physics self-explanations. *J. Learn. Sci.* 1, 69–105. doi: 10.1207/s15327809jls0101_4

Chiu, J. L., and Chi, M. T. H. (2014). "Supporting self-explanation in the classroom," in *Applying Science of Learning in Education – Infusing Psychological Science into the Curriculum*, eds V. A. Benassi, C. E. Overson, and C. M. Hakala (Washington, D.C: Society for the Teaching of Psychology), 91–103.

Conati, C., and Merten, C. (2007). Eye-tracking for user modeling in exploratory learning environments: an empirical evaluation. *Knowl. Based Syst.* 20, 557–574. doi: 10.1016/j.knosys.2007.04.010

Cook, S. W., and Goldin-Meadow, S. (2006). The role of gesture in learning. Do children use their hands to change their minds? *J. Cogn. Dev.* 7, 211–232. doi: 10.1207/s15327647jcd0702_4

Courant, R., and Robbins, H. (2010). *Was Ist Mathematik? [What is mathematics?] (5. Auflage)*. Berlin: Springer.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., and Paas, F. (2011). Improved effectiveness of cueing by self-explanations when learning from a complex animation. *Appl. Cogn. Psychol.* 25, 183–194. doi: 10.1002/acp.1661

Durkin, K., and Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learn. Instr.* 22, 206–214. doi: 10.1016/j.learninstruc.2011.11.001

Edwards, L. D. (2008). Gestures and conceptual integration in mathematical talk. *Educ. Stud. Math.* 70, 127–141. doi: 10.1007/s10649-008-9124-6

Edwards, L. D., and Robutti, O. (2014). "Embodiment, modalities, and mathematical affordances," in *Emerging Perspectives on Gesture and Embodiment in Mathematics*, eds L. D. Edwards, F. Ferrara, and D. Moore-Russo (Charlotte, CA: Information Age Publishing), 1–23.

Emmorey, K., and Casey, S. (2001). Gesture, thought and spatial language. *Gesture* 1, 35–50. doi: 10.1075/gest.1.1.04emm

Eppler, M. J. (2006). A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Inf. Vis.* 5, 202–210. doi: 10.1057/palgrave.ivs.9500131

Ericsson, K. A., and Simon, H. A. (1993). *Protocol Analysis*. Cambridge, MA: MIT Press.

Fiorella, L., and Mayer, R. E. (2016). Eight ways to promote generative learning. *Educ. Psychol. Rev.* 28, 717–741. doi: 10.1007/s10648-015-9348-9

Gallese, V., and Lakoff, G. (2005). The brain's concepts. The role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310

Gerofsky, S. (2010). Mathematical learning and gesture - Character viewpoint and observer viewpoint in students' gestured graphs of functions. *Gesture* 10, 321–343. doi: 10.1075/gest.10.2-3.10ger

Gerofsky, S. (2014). "). Making sense of the multiple meaning of "embodied mathematics learning," in *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, Vancouver, 145–152.

Goldin-Meadow, S. (2003). *Hearing Gesture. How Our Hands Help us Think*. Cambridge: Belknap Press of Harvard University Press.

Goldin-Meadow, S. (2010). Gesture's role in creating and learning language. *Enfance* 3, 239–255.

Goldin-Meadow, S., Kim, S., and Singer, M. (1999). What the teacher's hands tell the student's mind about math. *J. Educ. Psychol.* 91, 720–730. doi: 10.1037/0022-0663.91.4.720

Goldin-Meadow, S., and Singer, M. (2003). From Children's Hands to Adults' Ears: gesture's role in the learning process. *Dev. Psychol.* 39, 509–520. doi: 10.1037/0012-1649.39.3.509

Goodwin, C. (2003). "Pointing as situated practice," in *Pointing - Where Language, Culture and Cognition Meet*, ed. S. Kita (New Jersey: Lawrence Erlbaum Associates).

Greene, J. A., Robertson, J., and Costa, L.-J. (2011). "Assessing self-regulated learning using think-aloud protocol methods," in *Handbook of Self-Regulation of Learning and Performance*, eds D. H. Schunk, and B. Zimmerman (New York, NY: Routledge), 313–328.

Greiffenhagen, C., and Sharrock, W. (2005). "Gestures in the blackboard work of mathematics instruction," in *Proceedings of Interacting Bodies, 2nd Conference of the International Society for Gesture Studies*, Lyon.

Griffin, T. D., Wiley, J., and Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: concurrent processing and cue validity as constraints on metacomprehension accuracy. *Mem. Cogn.* 36, 93–103. doi: 10.3758/mc.36.1.93

Hankel, H. (1867). *Theorie der Complexen Zahlsysteme [Theory of complex number systems]*. Leipzig: Leopold Voss.

Hefter, M., Renkl, A., Rieß, W., Schmid, S., Fries, S., and Berthold, K. (2015). Effects of a training intervention to foster precursors of evaluativist epistemological understanding and intellectual values. *Learn. Instr.* 38, 11–22. doi: 10.1016/j.learninstruc.2015.05.002

Hegarty, M., Mayer, S., Kriz, S., and Keehner, M. (2005). The role of gestures in mental animation. *Spat. Cogn. Comput.* 5, 333–356. doi: 10.1207/s15427633scc0504_3

Hodds, M., Alcock, L., and Inglis, M. (2014). Self-explanation training improves proof comprehension. *J. Res. Math. Educ.* 45, 62–101. doi: 10.5951/jresematheduc.45.1.0062

Hostetter, A. B., and Alibali, M. W. (2007). Raise your hand if you're spatial: relations between verbal and spatial skills and gesture production. *Gesture* 7, 73–95. doi: 10.1075/gest.7.1.05hos

Hostetter, A. B., Alibali, M. W., and Kita, S. (2008). I see it in my hands' eye: representational gestures reflect conceptual demands. *Lang. Cogn. Process.* 22, 313–336. doi: 10.1080/01690960600632812

Hostetter, A. B., and Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture* 12, 62–83. doi: 10.1075/gest.12.1.04hos

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kenehan, G. (2007). "Examining function understanding from context using concept maps," in *Proceedings of the 29th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, eds T. Lamberg, and L. R. Wiest, Lake Tahoe, 93–99.

Kiewra, K. A. (1989). A review of notetaking: the encoding-storage paradigm and beyond. *Educ. Psychol. Rev.* 1, 147–172. doi: 10.1007/bf01326640

Kiewra, K. A., Benton, S. L., Kim, S. I., and Risch, N. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemp. Educ. Psychol.* 20, 172–187. doi: 10.1006/ceps.1995.1011

Kita, S., Alibali, M. W., and Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychol. Rev.* 124, 245–266. doi: 10.1037/rev0000059

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemp. Educ. Psychol.* 30, 242–262. doi: 10.1016/j.cedpsych.2004.10.001

Kobayashi, K. (2006). Combined effects of notetaking/reviewing on learning and the enhancement through interventions. *Educ. Psychol.* 26, 459–477. doi: 10.1080/01443410500342070

Koschmann, T., and LeBaron, C. (2002). Learner articulation as interactional achievement: studying the conversation of gesture. *Cogn. Instr.* 20, 249–282. doi: 10.1207/s1532690xci2002_4

Krause, C. M. (2016). *The Mathematics in Our Hands*. Wiesbaden: Springer Spektrum.

Krause, C. M., and Salle, A. (2016). "Learners' gestures when learning alone," in *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, eds C. Csíkos, A. Rausch, and J. Szitányi, Szeged: PME, 123–130.

Kurby, C. A., Magliano, J. P., Dandotkar, S., Woehrle, J., Gilliam, S., and McNamara, D. S. (2012). Changing how students process and comprehend texts with computer-based self-explanation training. *J. Educ. Comput. Res.* 47, 429–459. doi: 10.2190/ec.47.4.e

Lakoff, G., and Johnson, M. (1980). *Metaphors we Live*. Chicago, IL: The University of Chicago Press.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh*. New York, NY: Basic Books.

Lakoff, G., and Núñez, R. (2000). *Where Mathematics Comes From. How the Embodied Mind Brings Mathematics into Being*. New York, NY: Basic Books.

Lamnek, S. (2010).). *Inhaltsanalyse [Content analysis]. In Qualitative Sozialforschung (5. Auflage*. Weinheim: Beltz, 434–497.

Lin, L., and Atkinson, R. K. (2013). Enhancing learning from different visualizations by self-explanation prompts. *J. Educ. Comput. Res.* 49, 83–110. doi: 10.2190/ec.49.1.d

Luo, L., Kiewra, K. A., and Samuelson, L. (2016). Revising lecture notes: how revision, pauses, and partners affect note taking and achievement. *Instr. Sci.* 44, 45–67. doi: 10.1007/s11251-016-9370-4

Mayer, R. E. (2014). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Mayring, P. (2000). Qualitative content analysis. *Forum Q. Sozialforschung* 1. Available online at: https://www.qualitative-research.net/index.php/fqs/article/view/1089/2386

McEldoon, K. L., Durkin, K., and Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *Br. J. Educ. Psychol.* 83, 615–632. doi: 10.1111/j.2044-8279.2012.02083.x

McNamara, D. S. (2004). SERT: self-explanation reading training. *Discour. Process.* 38, 1–30. doi: 10.1207/s15326950dp3801_1

McNeill, D. (1992). *Hand and Mind - What Gestures Reveal About Thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2002). Gesture and language dialectic. *Acta Linguist. Hafniesia* 34, 7–37. doi: 10.1080/03740463.2002.10414607

McNeill, D. (2005). *Gestures and Thought*. Chicago: University of Chicago Press.

Melinger, A., and Kita, S. (2007). Conceptualisation load triggers gesture production. *Lang. Cog. Process.* 22, 473–500. doi: 10.1080/01690960600696916

Merten, C. (2002). *On the Use of Eye-Tracking in the Assessment of Self-Explanation in an Open Learning Environment*. Vancouver: The University of British Columbia.

Murayama, K., Pekrun, R., Lichtenfeld, S., and vom Hofe, R. (2012). Predicting long-term growth in students' mathematics achievement: the unique contributions of motivation and cognitive strategies. *Child Dev.* 84, 1475–1490. doi: 10.1111/cdev.12036

Nemirovsky, R., and Ferrara, F. (2008). Mathematical imagination and embodied cognition. *Educ. Stud. Math.* 70, 159–174. doi: 10.1007/s10649-008-9150-4

Neuman, Y., Leibowitz, L., and Schwarz, B. (2000). Patterns of verbal mediation during problem solving: a sequential analysis of self-explanation. *J. Exp. Educ.* 68, 197–213. doi: 10.1080/00220970009600092

Neuman, Y., and Schwarz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem-solving. *Br. J. Educ. Psychol.* 68, 15–24. doi: 10.1111/j.2044-8279.1998.tb01271.x

Núñez, R., Edwards, L. D., and Matos, J. F. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educ. Stud. Math.* 39, 45–65.

Núñez, R., and Freeman, W. J. (1999). *Reclaiming Cognition: The Primacy of Action, Intention and Emotion*. Thorverton: Imprint Academic.

Pika, S., Nicoladis, E., and Marentette, P. (2006). A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer? *Bilingualism* 9, 319–327. doi: 10.1017/s1366728906002665

Pirolli, P., and Recker, M. M. (1994). Learning strategies and transfer in the domain of programming. *Cogn. Instr.* 12, 235–275. doi: 10.1207/s1532690xci1203_2

Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educ. Stud. Math.* 70, 111–126. doi: 10.1007/s10649-008-9127-3

Radford, L., Edwards, L., and Arzarello, F. (2009). Introduction: beyond words. *Educ. Stud. Math.* 70, 91–95.

Recker, M. M., and Pirolli, P. (1995). Modeling individual differences in students'. Learning strategies. *J. Learn. Sci.* 4, 1–38. doi: 10.1207/s15327809jls0401_1

Renkl, A. (1997). Learning from worked-out examples: a study on individual differences. *Cogn. Sci.* 21, 1–29. doi: 10.1207/s15516709cog2101_1

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learn. Instr.* 12, 529–556. doi: 10.1016/s0959-4752(01)00030-5

Renkl, A. (2005). "The worked-out examples principle in multimedia learning," in *The Cambridge Handbook of Multimedia Learning*, ed. R. E. Mayer (Cambridge: Cambridge University Press), 229–245. doi: 10.1017/cbo9780511816819.016

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cogn. Sci.* 38, 1–37. doi: 10.1111/cogs.12086

Renkl, A., Schworm, S., and Hilbert, T. S. (2004). "Lernen aus Lösungsbeispielen: eine effektive, aber kaum genutzte Möglichkeit, Unterricht zu gestalten [Learning from exemplary solutions: an effective but overlooked instructional strategy]," in *Bildungsqualität von Schule*, eds M. Prenzel, and J. Doll (Münster: Waxmann), 77–92.

Renkl, A., Stark, R., Gruber, H., and Mandl, H. (1998). Learning from worked-out examples: the effects of example variability and elicited self-explanations. *Contemp. Educ. Psychol.* 23, 90–108. doi: 10.1006/ceps.1997.0959

Rittle-Johnson, B., Fyfe, E. R., Loehr, A. M., and Miller, M. R. (2015). Beyond numeracy in preschool: adding patterns to the equation. *Early Childh. Res. Q.* 31, 101–112. doi: 10.1016/j.ecresq.2015.01.005

Robutti, O. (2005). "Hearing gestures in modelling activities with the use of technology," in *Proceedings of the 7th International Conference on Technology in Mathematics Teaching*, Bristol: University of Bristol, 252–261.

Robutti, O. (2006). Motion, technology, gestures in interpreting graphs. *Int. J. Technol. Math. Educ.* 13, 117–125.

Roy, M., and Chi, M. T. H. (2005). "The Self-Explanation Principle in Multimedia-Learning," in *The Cambridge Handbook of Multimedia Learning*, ed. R. E. Mayer (Cambridge: Cambridge University Press), 272–286.

Sabena, C. (2008). "). On the semiotics of gestures," in *Semiotics in Mathematics Education*, eds L. Radford, G. Schubring, and F. Seeger (Rotterdam: Sense Publishing), 19–38. doi: 10.1163/9789087905972_003

Salle, A. (2014). "Learning with interactive animated worked-out examples in groups of two," in *Proceedings of the 38th Conference of the International Group*

for the Psychology of Mathematics Education, Vol. 5, eds P. Liljedahl, C. Nicol, S. Oesterle, and D. Allan, Vancouver: PME, 81–88.

Sassenberg, U., and van der Meer, E. (2010). Do we really gesture more when it is more difficult? *Cogn. Sci.* 34, 643–664. doi: 10.1111/j.1551-6709.2010.01101.x

Schwartz, D. L., and Black, J. B. (1996). Shuttling between depictive models and abstract rules: induction and fallback. *Cogn. Sci.* 20, 457–497. doi: 10.1207/s15516709cog2004_1

Schworm, S., and Renkl, A. (2006). Computer-supported example-based learning: when instructional explanations reduce self-explanations. *Comput. Educ.* 46, 426–445. doi: 10.1016/j.compedu.2004.08.011

Sfard, A. (2008). What's all the fuss about gestures? A commentary. *Educ. Stud. Math.* 70, 191–200. doi: 10.1007/s10649-008-9161-1

She, H. C., and Chen, Y.-Z. (2009). The impact of multimedia effect on science learning: evidence from eye movements. *Comput. Educ.* 53, 1297–1307. doi: 10.1016/j.compedu.2009.06.012

Stark, R. (1999). *Lernen mit Lösungsbeispielen: Einfluss Unvollständiger Lösungsbeispiele auf Beispielelaboration, Lernerfolg und Motivation [Learning with Exemplary Solutions: Incomplete Examples and Their Influence on Elaboration, Motivation and Success]*. Göttingen: Hogrefe.

Streeck, J., and Kallmeyer, W. (2001). Interaction by inscription. *J. Pragmat.* 33, 465–490. doi: 10.1016/s0378-2166(99)00126-5

Swing, S., and Peterson, P. (1988). Elaborative and integrative thought processes in mathematics learning. *J. Educ. Psychol.* 80, 54–66. doi: 10.1037/0022-0663.80.1.54

van der Meij, J., and de Jong, T. (2011). The effects of directive self-explanation prompts to support active processing of multiple representations in a simulation-based learning environment. *J. Comput. Assist. Learn.* 27, 411–423. doi: 10.1111/j.1365-2729.2011.00411.x

Wagner Cook, S., Mitchell, Z., and Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition* 106, 1047–1058. doi: 10.1016/j.cognition.2007.04.010

Wartenburger, I., Kühn, E., Sassenberg, U., Foth, M., Franz, E., and van der Meer, E. (2010). On the relationship between fluid intelligence, gesture production, and brain structure. *Intelligence* 38, 193–201. doi: 10.1016/j.intell.2009.11.001

Wilkin, B. (1997). "Learning from explanations: diagrams can "inhibit" the self-explanation effect," in *Reasoning with Diagrammatic Representations II: Papers From the AAAI Fall Symposium*, ed. M. Anderson (Menlo Park, CA: American Association for Artificial Intelligence), 136–143.

Wirtz, M., and Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität [Rater Consistency and rater reliability]*. Göttingen: Hogrefe.

Wong, R. M. F., Lawson, M. J., and Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learn. Instr.* 12, 233–262. doi: 10.1016/s0959-4752(01)00027-5

Wylie, R., and Chi, M. T. H. (2014). "The self-explanation principle in multimedia learning," in *The Cambridge Handbook of Multimedia Learning (2. Auflage*, ed. R. E. Mayer (New York, NY: Cambridge University Press), 413–432. doi: 10.1017/cbo9781139547369.021

Yammiyavar, P., Clemmensen, T., and Kumar, J. (2007). "Analyzing non-verbal cues in usability evaluation tests," in *Usability and Internationalization. HCI and Culture. UI-HCII 2007* Vol. 4559 (Berlin: Springer), 462–471. doi: 10.1007/978-3-540-73287-7_55

Yeo, A., Ledesma, I., Nathan, M. J., Alibali, M. W., and Breckinridge Church, R. (2017). Teachers' gestures and students' learning sometimes "hands off" is better. *Cogn. Res.* 2, 41–52.

Yoon, C., Thomas, M. O. J., and Dreyfus, T. (2011). Grounded blends and mathematical gesture spaces: developing mathematical understandings via gestures. *Educ. Stud. Math.* 78, 371–393. doi: 10.1007/s10649-011-9329-y