Supporting Inquiry Learning: an Intellectual Mirror that Describes what It "Sees"

Michal Yerushalmy[1] · Shai Olsher[1] · Raz Harel[1] · Daniel Chazan[2]

Abstract

Much important research on the learning of mathematics with technology-supported inquiry has been devoted to learning with multiple-linked representations (MLR) as a mode of feedback. Like a mirror, MLR feedback helps students see their actions in one representation "reflected" in another. Yet, research has followed learning episodes where MLR feedback did not lead to concept formation and the achievement of curricular goals. This article reports on the potential of what might be thought of as a mirror that speaks. In response to example-eliciting tasks, students use interactive diagrams to create examples to which mathematical descriptions are automatically associated. Such descriptions may be thought of as another kind of linked mathematical representation system. Transitions feature in two ways in our analysis of students' use of this representation. At the level of student activity, we examine when students move between attending to textual descriptions and to the graphs that they describe. We are also interested in how attention to these descriptions and coordination with their own use of these words can support students in making a transition in their thinking from considering distance as only total distance traveled, to a co-ordinated view of distance including both total distance traveled and distance from a starting point. This article focuses on two example-eliciting motion tasks and two sets of descriptive words. We found that these sets of words helped students, while and after they were working with the diagram, to distinguish between total distance traveled and position with respect to a starting point.

Keywords Inquiry learning · Feedback · Student reflection · Technological supports for learning · Example eliciting tasks

✉ Michal Yerushalmy
michalyr@edu.haifa.ac.il

1 University of Haifa, Haifa, Israel

2 University of Maryland, College Park, MD, USA

This article reports on results from a design research effort in which students were interviewed while using an interactive diagram (Naftaliev & Yerushalmy, 2017) to generate position-over-time graphs to fit a given story about a bike trip. Students used a new kind of feedback on their work: technology-provided automatic characterization of their examples using mathematical terminology (for an elaboration of this strategy, see Yerushalmy et al., in press). With this feedback, students verified that they had created multiple examples that fit the story and that differed one from another. Students worked on two tasks with this diagram within the STEP (seeing the entire picture) platform which supports automatic characterization of students' examples using words (Olsher et al., 2016). This article focuses on feedback processes in which students interact with technological tools meant to support them in distinguishing *accumulated distance traveled* from *distance from starting point (*or *position)*. These tools are used to examine the range of examples students think of as trips involving a fixed total distance traveled.

The two tasks used in this study are example-eliciting tasks (Yerushalmy, 2020), that is, tasks in which students are asked to generate examples. We value such tasks because generating and verifying examples of a particular mathematical concept serves as an indicator of learners' understanding (Zaslavsky & Zodik, 2014). However, the tasks used in this article require that students do more than submit a single example; students are asked to submit multiple examples. Seeing the range of examples submitted that are as different as possible one from another provides another sort of insight into students' understanding of a concept. Building on the construct *personal example space* introduced by John Mason and colleagues in the context of conceptual understanding (Goldenberg & Mason, 2008; Mason, 2008), insight into a personal example space from our task comes from the variety of graphs representing their model of the motion situation described in the presented task.

Stimulated by the seminal work of Rina Hershkowitz (1990) on conceptual understanding, as configured for this task, the software environment provides students with feedback both on *critical characteristics* that indicate that a graph exemplifies the given story and on *non-critical characteristics* which indicate how graphs that exemplify a story may differ one from another. By design, students received feedback at four stages during their work on this activity, getting information about whether their examples meet the critical characteristics of an example as they are working on a first task. They then receive further information upon the conclusion of their work on this task and its submission: such feedback tells them both whether their examples fit the story and how they differ on non-critical characteristics of the trip. Finally, they receive similar in-task feedback on the second task and similar post-submission reports after the completion of the second task.

Transitions feature in this study in two ways. First, we use the transition to describe how the range of graphs students create to exemplify a story changes as students' source-path-goal schema evolves. We are interested in the transition from thinking about distance as only total distance traveled to a co-ordinated view of distance including both total distance traveled and distance from a starting point. We take such a transition in thinking when receiving feedback as an indication of the utility of that feedback.

A second way transitions feature in this study is at the level of student activity where we examine when students attend to different standard mathematical representation systems (that are linked dynamically with the technology). With the interactive diagram used in this article, as students work to create a graph to match a given story that is described to them in words, there are two dynamically linked representations: a graph of the relationship between the quantities specified by the words on the axes of a Cartesian plane and words that capture whether or not the graph meets the four characteristics of the story presented in the task.

When students have submitted their work, there is an additional set of descriptive words linked with their submitted graphs. These words connect the graph to the story by narrating aspects of the stories that the graphs would represent. In contrast to the four critical characteristics that indicate that a graph does exemplify the story, these non-characteristics are designed to help students think about the differences among the graphs they have submitted.

The article introduces our effort in six sections. The first section examines the literature on feedback and inquiry, with a focus on computer-provided feedback and multiple-linked representations (MLR). It introduces metaphors used in the literature on technology and feedback processes, metaphors that describe the interaction between technology and students during the inquiry. This literature suggests that it is useful to consider students' interactions with information supplied by technology as a potential meeting place of curricular-expected ways of speaking and ways of speaking that students bring to the classroom. At the same time, it is also useful to analyze students' interaction with the information provided by technology to learn how students use that information to further their own learning.

The second section lays out the research questions specific to the instructional context of the study that is examined in this study, while the third one details the methods used in this study. It explains how the data was collected, how the data was analyzed, and its presentation as a single vignette that captures key themes in students' interaction with characterizations of their work. The fourth section provides readers with a window into the data by presenting (as one vignette) key episodes where the students interacting with STEP learned to use the feedback received from the platform to support their learning in completing the tasks that they were assigned.

The fifth section summarizes what we learned about the role of technology-provided automatic feedback in supporting students' learning during the inquiry in the context of motion and understanding of distance. The final section closes by moving beyond the research questions specific to this curricular context and reflecting on the potential and drawbacks for supporting student reflective inquiry by automatically associating textual descriptions with students' submissions.

Inquiry Learning and Feedback Processes: Exploring a New Strategy

There are many kinds of activity structures that occur in mathematics classrooms and many goals that teachers have for classroom activities. One such educational goal that plays a central role in inquiry learning is to have students reflect on their

own ideas. This section embeds this article—and its aim to support students' reflection on their own thinking—in the literature on feedback processes, particularly on the centrality of communication with words in cognitive development.

Arranging Feedback Processes as Part of Mathematics Learning

We start by examining the purpose of feedback processes occurring as part of learning mathematics, specifically when learning with technology. In the literature on learning with technology, the term "feedback" is often used to describe the information that technology presents to students or their teachers regarding aspects of a learner's performance or understanding. Learner's performance aspects may include corrective information, an alternative strategy, information to clarify ideas, encouragement, or simply the correct answer. Hattie and Timperley (2007) conceptualized *feedback* to be "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (p. 81).

Carless (2015) challenges this definition of feedback for only describing a one-way transfer of information from an agent to the learner, stating that, "Feedback is a dialogic process in which learners make sense of information from varied sources and use it to enhance the quality of their work or learning strategies" (p. 192). By way of contrast, this view positions learners as active agents, making sense of information and using it to enhance their work. Thus, in the literature we cite, the term *feedback* is used both to describe processes in the classroom and to identify the information that technology can present for use as part of such processes. Although we conceptualize feedback as a process, to be true to our citations, we use "feedback" in these two different ways.

A common practice in the study of feedback processes is to examine whether, and how, feedback helps students close the gap between their current and expected performance. Yet, studies agree that the perceived effectiveness of feedback is highly contextual. In studying the effectiveness of feedback, researchers have distinguished between numerous types of formative feedback along a variety of dimensions, often distinguishing between the main types of verification and elaborated information. There have been researching and development attempts of technology-based, textual-elaborated reports to students.

These were most often verbal reports provided, with or without technology, aimed at judging and supporting the direct acquisition of a concept: usually, they are targeted at the right answer and procedures. Other systems seek to simulate a mathematical conversation using prepared hints or comments that address the performance of the learner compared with the envisioned or expected performance. While simple verification of correctness is at times considered to be more effective than elaborated feedback, which explains why a response is correct or incorrect, further research conclusions have led to opposite claims: for example, van der Kleij et al. (2015) review of on-line personal feedback concluded that elaborated information was more effective.

Other studies of feedback concern the effects of immediate as opposed to delayed feedback on learning outcomes. Among the positive effects of immediate feedback, studies indicate helping students in their decision or motivation to practice the tasks and providing an explicit association between outcomes and causes during problem solving (Nathan et al., 2005; Shute, 2008). A negative effect of immediate feedback may be that it leads to dependence on information that is not available during transfer tasks, and it may lead to less care in the choice of answers and may impede metacognitive activities (Nathan & Koedinger, 2005). On the positive side, delayed feedback may encourage learners' engagement in active cognitive and metacognitive processing, creating a sense of autonomy and self-regulation (Bokhove & Drijvers, 2012; Shute, 2008).

Metaphors for Interaction Between Technology and Students During Inquiry

A recent meta-review (Jensen et al., 2021) of the literature on feedback identified two metaphors that are appropriate for supporting students' reflection on their own thinking: tool metaphors and dialogic metaphors. Tool metaphors are highlighted in the work of Carless and Boud (2018), who suggest that a tool metaphor "highlights learner agency and the learner's capacity to use the tool" (p. 7).

One example of a technology that has been thought of as such a tool for providing feedback to help students reflect on their own thinking is that of the dynamic MLRs. Traditional dynamic MLRs reflect back the actions of the user by reproducing user-constructed mathematics in a different representation and thus they see their own ideas differently. For example, MLR feedback allows one to check whether geometrical constructions are based on mathematical considerations and not simply accomplished visually (Laborde & Laborde, 2014). The metaphorical term "intellectual mirror," coined by Schwartz (1989), articulates the essence of the contribution of this strategy for supporting the technology-based inquiry. Schwartz referred to mirroring in which feedback reflects a user's actions back to them, without judgment, just as a mirror reflects an image. The aim of such mirroring is to support a process of self-reflection.

Note that, in this strategy of dynamically linked representations, users transition between the examination of a phenomenon and its representation and between different representations of the same phenomenon. In such contexts, users must learn to integrate the information available to them from such transitioning and use that information as a tool for learning.

By contrast to viewing feedback as information, using dialogic metaphors, feedback is seen as a cyclical process that serves social purposes. In such processes, participants in the process (including artifacts and other agents) share both agency and responsibility for creating a productive and meaningful feedback process that will support reflection. This sort of metaphor is the basis for emphasizing the importance of examining dialog between the provider and the receiver of the feedback, even when that provider is a machine.

A New Strategy: Characterizing Mathematical Examples in Words

We now turn to the particular strategy for providing students with information to reflect upon their own ideas that are the focus of this design research effort: giving students mathematical descriptions in words of their submitted examples. We conceptualize the sorts of textual descriptions of examples that are the focus of this article as an alternative to information that is more commonly provided as feedback: evaluation of the correctness of responses. In this sub-section, building on earlier research on feedback and on the nature of the learning of mathematics, we examine two reasons that characterizing students' examples using words might be a valuable aspect of feedback processes aimed at supporting student reflection. First, words that capture characteristics of sets of examples can capture succinctly important variations between many examples. Second, in the context of example-eliciting tasks, though students may have to come to learn terminology with which they are unfamiliar, using the vocabulary that teachers have a curricular responsibility to teach in feedback can help students connect that vocabulary to examples of the phenomena being studied.

The use of text in parallel with graphical representations has been explored as a support for learning that might offer students ways to understand relations and structures conceptually. For example, Talmon and Yerushalmy (2004) describe a design that sought to link visual feedback offered by the drag mode in a geometry construction program to an automatically-created list of words describing the procedure students used to make a geometrical construction. Each procedure captured as a list of words in this environment can be enacted on a variety of initial shapes. Thus, by varying the starting points, there are many diagrams that can be created by each captured procedure. A set of words can describe a large example space.

Based on a similar set of considerations about connections between words and examples, Schwartz and Yerushalmy (1995) connect words and images in a different way. They offer a mathematical bridging language between the functional relationships described in a story or a simulation and graphical representations of relationships between quantities. They argue that the bridging language that they have created can be used to characterize quantities and the rates of change of quantities in a succinct way and thus provide language that supports problem-solvers in modeling a wide variety of situations.

However, such reliance on words has challenges as a strategy for supporting reflection and learning. The words used in feedback to students may provide descriptions in a lexicon of objects and actions in ways that may be different from how students would use these same words. Students must then negotiate conflicts between the words provided by STEP and their own uses of these words.

Sfard (2007) anticipates this sort of challenge and argues that.

> Most of the time, our discourses remain consistent with our experience of reality […] Without other people's example, children may have no incentive for changing their discursive ways. From the children's point of view, the discourse in which they are fluent does not seem to have any particular weaknesses as a tool for making sense of the world around them. (p. 574)

Thus, for Sfard, the challenges in using words as feedback are ones that should be expected as part of a learning process and may actually provide resources for reflection. Her perspective suggests that we pay attention to how a discursive change could occur for a learner, through which the learner could become aware of new possibilities and arrive at a new vision of the mathematical landscape they are exploring.

From this perspective, characterizing students' examples in words might be thought of as providing a mirror that speaks, in some ways like the mirror in the Snow-White tale. In that story, a character asks a magic mirror a question: "Who is the fairest one of all?" Usually, the mirror answers the character by saying that they are the fairest. But, at one point in the story, the magic mirror unexpectedly responds "snow white."

In this article, we describe a mirror that describes an instance, rather than simply reflecting back an image or answering an evaluative question. The sort of "talking mirror" we aim to create "speaks" to the user by offering a task designer's description of a learner's example.

A mirror that describes in words the characteristics that make an example an example of a particular concept (rather than simply saying that it is an example) can be understood both in terms of a tool metaphor and a dialogic metaphor. As a tool, the provided information is designed to reflect a user's actions back to them and can be used as a tool. At the same time, this feedback on examples is dialogical, in that it is designed to support a learner's interaction with descriptions of characteristics of examples that are relevant to particular learning goals, and which are articulated in the vocabulary of the designer. Thus, this information offers resources to a student for transitioning to a new way of thinking: for example, in our case, distinguishing total distance travel from the position. In this context, feedback processes can include attempts to resolve conflicts initiated by the presentation of mathematical information in words.

In the work we are doing currently with the STEP platform, the words—the mathematical characteristics used to describe students' work—are chosen by designers and educators. Students must learn through experiments to appreciate the information provided to them as a resource by software designers and educators. Although we study whether a designed collection of words may act similarly to the contribution to classroom discourse brought into a discussion by teachers to challenge students' ideas and conclusions (Chazan & Ball, 1995), personal feedback in this article is not the means for individualized learning or a substitute for teacher feedback as assessment. Instead, it is part of the student's interaction, where artifacts designed as tools by others may become instruments for learning.

Related to automatically provided feedback for supporting student reflection, our overarching research question in this article is whether the characterization of students' examples created automatically by STEP serves as feedback, which enables students to reflect upon their own understandings of the examples they have created.

Research Questions Specific to the Instructional Context of the Study

The curricular context of this article involves the mathematization of descriptions of position over time of moving objects and what Lakoff and Nunez (2000) refer to as the Source–Path–Goal schema which derives from physical experiences such as motion along a path. This schema consists of three components: initial state (= source position), final state (= goal position), and action sequence (= movement from the source to the goal). However, the limited experiences students have with this schema hamper student inquiry. We seek to deepen students' engagement with mathematizing a situation involving the Source–Path–Goal schema in two challenging ways. First, we seek to help students work to understand the relative positioning of reference points on a Cartesian graph, for example, the importance of the conceptualization of a reference point or the origin in mathematical representations (Radford, 2009). Radford argues that, without making a shift from local to abstract space and conceptualizing the idea of relative places on the graph system, students would not be able to understand the Cartesian graphs appearing in their textbooks.

Second, we seek to develop students' awareness of the situational structure of motion in the context of the possibility for changing directions of travel. In particular, our work focuses on helping students understand that riding towards a target specified by distance and time of an ending point is not identical to thinking about a constantly increasing graph of position over time, but instead can include a graph that has sections that decrease and correlate with riding back along the path towards the starting point. The challenge is to realize that, when one changes direction, as one continues to accumulate total distance traveled, one's graph of distance from the starting point can be a decreasing graph.

In describing the challenge of developing this understanding, Nemirovsky (1994) describes Laura's challenge as occurring when the mathematical representation of *more* does not correspond to her image of *more*, meaning that the graph must show an increase in quantity. Yerushalmy and Gilead (1999) demonstrate a few occurrences of stories involving motion in a round trip and suggest that the functional considerations of a negative slope, which is not considered as an obstacle when approached symbolically, introduce a challenge requiring the shift in the schema of rate from describing speed to velocity and the inclusion of issues of the direction of the motion (Gilead & Yerushalmy, 2006). Their evidence shows that most students would sketch the situational model according to what they consider as the chronological stages of the trip, in comparison to the percentage of students who were able to envision an alternative structure that will lead to an equivalent but easier one to solve the problem. Schnepp and Chazan (2004) discussed a related issue with the distinction between speed and velocity.

As students' ideas grow and change, in this curricular context, we examine when students attend to different standard mathematical representation systems

(that are linked dynamically with the technology). As students work to create a graph to match a given story that is described to them in words, there are two dynamically linked representations: a graph of the relationship between the quantities specified by the words on the axes of a Cartesian plane and words that capture whether or not the graph meets the four characteristics of the story presented in the task that mean that the graph represents the relationship between quantities that could be described by the words in the story. When students have submitted their work, there is an additional set of descriptive words linked with their submitted graphs. These words connect the graph to the story by narrating aspects of the stories that the graphs would represent. In contrast to the four characteristics that indicate that a graph does exemplify the story, these characteristics are designed to help students think about the differences between the graphs they have submitted.

The goal of this study is to examine how students navigate transitions between representations—between words used to describe a situation (as one representation), in this case of motion, and mathematical objects and relationships as represented in a graph, as a standard kind of mathematical representation. Specifically, in the context of automatic characterization of student examples in words, we examine three research questions during two stages of student inquiry: interaction with the inquiry task, and interaction with a post-submission report, both of which include characterizations of student examples in words. At each stage, we examine three research questions:

1. In what ways can feedback that characterizes student-created examples with critical and non-critical characteristics help students co-ordinate between words that capture the situation and graphical representations of that situation that students have created, hence transitioning in their thinking?
2. For what purposes might students utilize co-ordinated graphs and characterization of those graphs in their work in the STEP platform?
3. Is there evidence that, as a result of interaction with the automatic characterization of students' examples, students' personal example spaces change?

Methodology

The study reported in this article is part of a research project entitled "Example-based online assessment of mathematical reasoning: affordances of personal elaborated feedback," funded by the Israeli Science Foundation. The project focused on designing ways of reporting information to students and pilot studies exploring ways for engaging students with the information more effectively. To prepare for the presentation of data in the next section, in this section we begin by describing how the STEP platform is designed to provide students with feedback that describes examples they create back to them and how the capacities of STEP were used in the context of a particular learning task.

324 Digital Experiences in Mathematics Education (2023) 9:315–342

Research Tools: the Platform and the Task

In this study, we use STEP (Olsher et al., 2016), an online platform designed to support teachers' work in assessing various open-ended, example-eliciting tasks (EETs), and to support technology-enhanced didactical situations that involve learning processes with non-judgmental individual feedback. The tasks include interactive diagrams (applets) in *GeoGebra* (Hohenwarter et al., 2009).

With the importance of example generation in mind, many tasks in STEP are example-eliciting (Yerushalmy, 2020). The tasks we are reporting on were designed with an illustrating interactive diagram (Naftaliev & Yerushalmy, 2017) as a digital learning environment through which students submit examples. The illustrating interactive diagram usually consists of a single representation, and it does not offer links, in the traditional MLR sense, between representations. Thus, it offers fewer opportunities for engaging students in inquiry, but offers the simplicity of use; in the tasks presented in this article, students can construct graphs by direct manipulation (dragging) of a given sketch. Students were asked to construct distance traveled-from-a-starting-point-over-elapsed-time graphs. Each task required students to construct and submit three different motion graphs that meet given conditions.

Students were encouraged to explore the interactive diagram before they decided which states of the diagram they would like to include as examples in their submission. They were allowed to reconstruct and change their decisions before submitting.
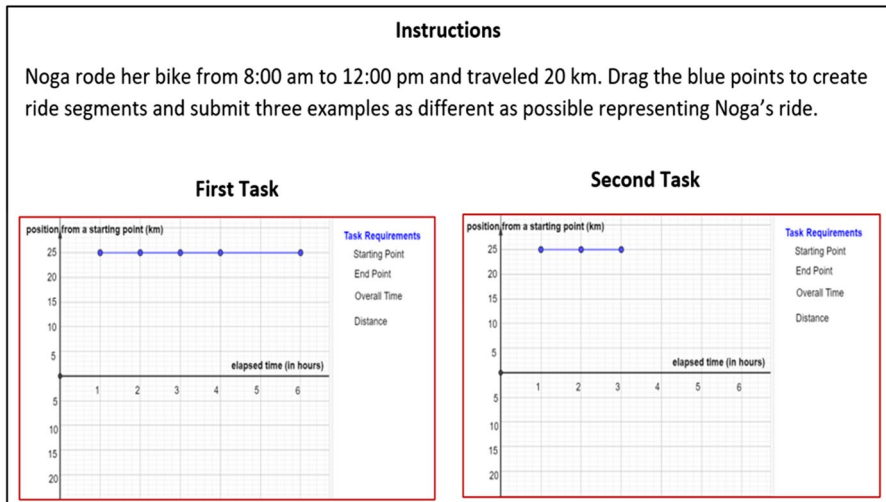
Fig. 1 Two tasks related to a single situation involving a bicycle trip (the tasks differ in the number of segments in the interactive diagram out of which students can construct a trip)
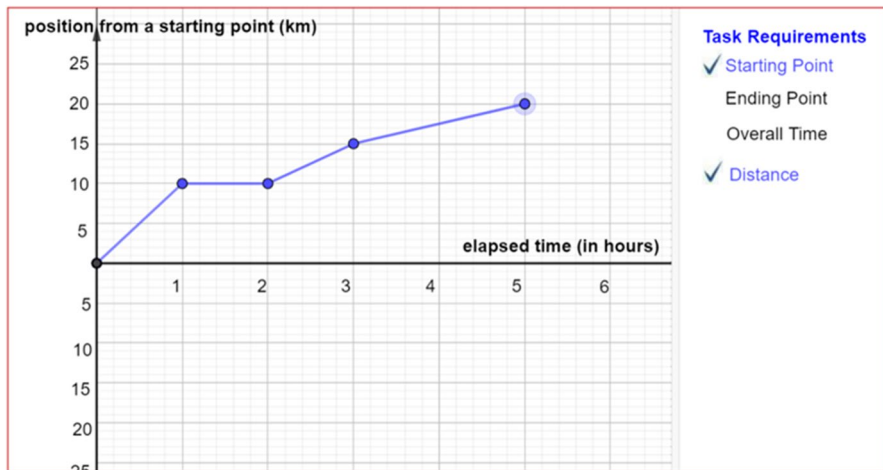
Fig. 2 Interactively updated information about meeting task requirements

Pre-submission Automatic Linked Descriptions of Student Work

Based upon the principles articulated by Harel et al. (2022), we sought to design a set of words that STEP could use to describe students' graphs in relation to the situation of a bike ride. As shown on the right-hand sides of the two screens captured in Fig. 1 above, and in Fig. 2 below, this *GeoGebra* interactive diagram includes a list of requirements that examples of a Noga-trip should meet.

There were four such conditions. Submitted examples were supposed to show:

- an overall travel time of 4 h (overall time);
- a total distance covered of 20 km (distance);
- a starting point of (0, 0) (starting point);
- an end point where $x = 4$ (end point).

As students work on an example, whenever an element of the constructed graph met a requirement, *GeoGebra* highlights the relevant text on the list.

Getting four check-marks tells students that their graph is an example that matches the given verbal description. In that sense, the task can be seen as defining a construct that one might call a "Noga-trip" and these four characteristics (task requirements) are the critical characteristics, in the sense of Hershkowitz (1990), for defining this construct.

Post-submission Automatic Linked Descriptions of Student Work

In addition to the information provided to students as they explored to help them identify whether their graphs represented Noga-trips, students' submissions were stored and automatically analyzed by STEP to produce post-submission individual reports for the students (hereinafter, called 'a post-submission report') (Olsher

et al., 2016). For all of the examples submitted by students, the post-submission report for these tasks was produced on the list of task requirements that appeared during exploration, as well as on a list of additional mathematical characteristics of the submissions, non-critical characteristics of this construct in Hershkowitz's (1990) terms. Having these additional characteristics available as feedback was intended to provide students with support for understanding their personal example space of a Noga-trip.

More generally, part of the task design process in STEP includes a priori definitions of mathematical characteristics of student examples that can provide useful feedback. As part of authoring a task, designers provide the platform direction about the characteristics of student submissions to be checked and associated with examples. These can then be presented to students as they work in the diagram or upon submission of their work.

With these particular motion tasks, we considered characteristics regarding three aspects which were defined as central to the goals of the work on the task, understanding of *speed*, *direction*, and *position*. Under *position* are requirements that should be met regarding the starting- and end-points, total distance and total time, as well as characteristics regarding revisiting the starting point or ending the trip at that point. The characteristics of student work related to *speed* indicate constant speed, changing speed (along the graph), or a rest or stop, but without indicating direction. *Direction* can either be changing or not. Note that we decided not to speak in terms of velocity, but rather to include information both about speed and about direction so that students would need to co-ordinate the information about speed and direction. To summarize for this activity, we developed the following list of characteristics (Fig. 3).

Critical characteristics that define a graph as being an example of a representation of a Noga-trip	Overall time meets task requirements Overall distance meets task requirements Starting point meets task requirements End point meets task requirements
Non-critical characteristics of Noga-trip (characteristics that distinguish among Noga-trips and that can be true for Noga-trips and for non-Noga-trips)	Noga stopped at least once Noga rode at different speeds Noga changed direction at least once Noga finished riding in the city from which she left Noga passed through the city from which she left

Fig. 3 Reported characteristics for the two tasks on Noga's bike trip

The role of the characteristics in Fig. 3 is both to articulate mathematical ideas in the context of the situation and to introduce a discourse about the phenomena that the students can interact with, which they can either reject or use to refine their submissions. These characteristics can provide information that goes beyond whether students' submissions are right or wrong. Critical characteristics provide information on whether or not an example is an example of a Noga bike trip; non-critical characteristics can give students a sense of the breadth of their personal example space. Taken together, these two kinds of characteristics have the potential to challenge the students' current perspectives and may be resources to create a shift in students' understandings.

Figure 4 illustrates how multiple examples submitted by students are characterized by the different sets of words: a limited set of critical characteristics, as well as a wider range of possible non-critical characteristics that can go beyond the scope of the situation.

Participants and Procedure

The participants in this study were four secondary school students from different schools, learning geometry, algebra, and numerical thinking according to the national curriculum. The participants focus in this stage of their schooling on

Fig. 4 Post-submission report for work on Task 1

fundamental skills and connections between the different topics and do not specifically prepare for the matriculation examinations that will take place in two parts: one in the 11th grade and the other in the 12th grade. The students are acquainted with graphical representations of functions, as well as with word problems in general and, specifically, motion problems. The use of a Cartesian plane to represent a motion problem has the potential to be a novelty for them.

Each student carried out an activity consisting of a task-based interview conducted by the third author. The role of the interviewer was to motivate the participants to express their thoughts and to help with technical matters. The activity contained two similar example-eliciting tasks. Each one required students to submit three examples demonstrating different Noga-trips. Each Noga-trip in the first task was supposed to be constructed using four draggable graph segments, while in the second task each trip was constructed of two graph segments. The second task, therefore, enabled a narrower space of potential trips. Once the students submitted their examples for the first task, they had the opportunity to interact with the information provided by the post-submission report. They then proceeded to engage with the second task and went through a similar procedure.

Data Analysis

When examining the responses of individual students, STEP can help begin to characterize what Mason (2008) calls a personal example space. As articulated earlier, two types of words were used to characterize the mathematics of the situational model which are analyzed automatically by STEP when students submit a graph. We treat these types of words as providing students with information about critical characteristics which indicate that a graph exemplifies a story and non-critical characteristics which do not determine whether or not a graph exemplifies a story, but indicate how graphs that exemplify a story may differ from one another. In the tasks in this article, both kinds of characteristics are given to the student: critical characteristics as they work on their submissions and non-critical characteristics after they have submitted their work.

Using the resources described in Fig. 5 as analytical constructs, we started by analyzing student submissions for critical and non-critical characteristics. The initial stage of data analysis included identifying which characteristics appeared and which did not in the submitted examples. Next, student interactions with the tasks with the immediate pre-submission information and with the post-task reports were separately segmented from the video interview recordings. Each interaction was coded into a specific segment.

In the following stage, data from each interaction was arranged based on which resource of the task was involved:

- the words that captured the situation (by the given story in the task);
- the standard mathematical representations (the graph system);
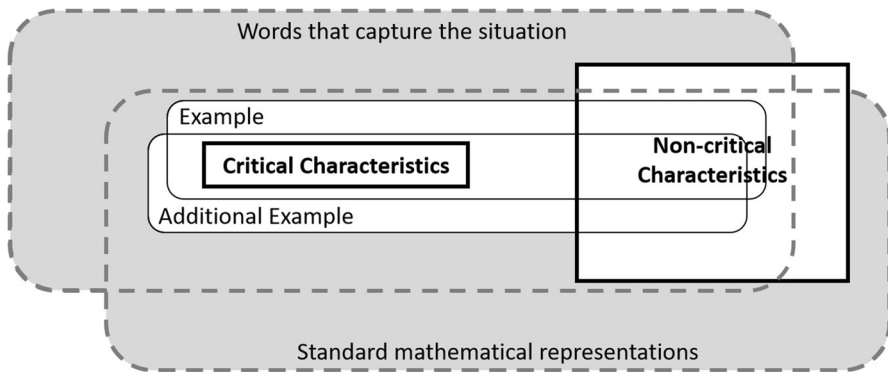
Fig. 5 Linked resources available to students when working on the Noga-trip tasks

- the characterizations of student work in words, in terms of the critical or the non-critical characteristics.

The data was then used to identify and code the focus of interactions: situations characterized by transitions between representations and situations that focus on transitions in the meaning of distance.

The coding of the two types of situations was then used to define three categories of students' attention: attention to different representations within a single example; attention to differences and similarities across examples; attention to the meaning of distance. These categories were then used to illustrate students co-ordinating between the two representations (research question 1), their purposes in utilizing the co-ordinated characterization of the situation and the graph (research question 2), and demonstrating changes in students' personal example spaces (research question 3).

Presentation as a Single Vignette

At this early stage in our research effort, in the results section, rather than focusing on how individual students differed in their use of the tool, we seek to communicate the range of ways in which the students piloting these tasks used the tool. Thus, rather than presenting four different case studies, each indicating the ways in which a particular student used the tool, we present what transpired in the interviews in the form of a single vignette about a single fictional student, named Sarah. The words attributed to Sarah are ones that were spoken by one of the four interviewees.

In this way, we are able to illustrate the major categories of interactions between students and STEP feedback that emerged from the coding of the different interviews. This presentational decision supports a coherent flow of interactions with the task and the different resources that enables the reader to focus on the significant phenomena that arose during different interviews with different students, while following an imaginary single student's problem-solving process for the relevant task.

Presentation of data: Representing Student Interactions with Descriptions of their Graphs

In this section, we present the ways that interaction with the characterization of student graphs involves students co-ordinating between the words that relate to the situation and the graphs they have submitted. This vignette illustrates how in their work students transitioned between the different resources provided by STEP. As students work with these different resources, we are then able to identify transitions in their thinking about distance.

Episode 1: Working on a Single Example of a Noga-Trip

The first episode focuses on transitions among words that capture the situation and student-created graphs. This transition is taking place while the student is focusing on the critical characteristics of a Noga-trip when creating a graph.

Sarah read the task instructions (Fig. 1) and started by reflecting out loud on the components of the given interactive diagram [spoken turns 1–3 reported below]. First, she pointed to a conflict between her reading of the story and the graphical representation regarding the requirement of the start time of motion [turn 2]. The start time was 8 o'clock, which she could not identify on the graph in the interactive diagram.

1. Okay. It's written that she left at 8:00 in the morning until 12:00 and traveled 20 km … So, I need to place these [points to the blue dots and intervals].
2. Wait … but I don't have eight on the axes …
3. Maybe I'll try to move the points a bit and something will happen [moves the left-most point and looks for changes in the list]. Mm … Nothing is changing.
4. Well, I'll have to think harder [reads the first property in the list] "The starting point" is checked so it meets the conditions of the task … ah.
   So, I'll position that point [Sarah starts dragging the left point] at the beginning of the ride.
5. Starting point … This axis is the time that has passed and this is the distance from the starting point. Oh, so that's clear. It should be at zero [Sarah places the left dot at the origin and the top characteristic is marked as blue].

Realizing the functionality of the words that were already checked, Sarah continues to construct the graph:

6. Ha! So that's what it does! It lights up when it's right. What a beauty. I made some progress.
7. Now I also understand the axes. So, I'll build a graph so that 4 h will pass and the road traveled will be 20 km. And this [indicates the condition Ending point meets requirements] will then be marked.

Sarah positioned the second point, the third, the fourth, and then the fifth, and created the graph shown in Fig. 6. All four conditions characterizing the requirements were then checked. Sarah started by placing the initial point at (0, 0) [turn 5], then proceeded according to the order of the events in time to meet the requirement of 4 h and 20 km [turn 7].

Initially, in this episode, Sarah did not see how she could solve the task of creating a Noga-trip. By dragging and watching for any indication of the change in the words describing the requirements, she realized that the information that automatically appeared in her diagram represented the requirements of a Noga-trip [turns 4–5].

Using the metaphor of the mirror to describe the instrumentation of the list of required characteristics, as Sarah dragged the graph, she used the information displayed in the mirror to guide her actions, perhaps in the way that a driver uses a mirror when backing into a tight spot. The information provided in the mirror guided her action and validated a course of action. Having figured out how the information guided her, she then used the way it worked to plan a course of action and to imagine what the result would look like in the mirror.

Episode 2: Creating Graphs of Multiple Noga-Trips

In this episode, Sarah is searching for ways to construct other graphs that are examples of a Noga-trip. That leads her to make distinctions among elements in the graph.

Satisfied with identifying the terms provided in the feedback and meeting the requirements for one Noga-trip, Sarah turned to analyze the construction process that she undertook for the first example in order to reflect on how to meet the requirement of submitting a variety of examples. She had to construct at least three different examples, which led her to generalize the actions she had carried out so
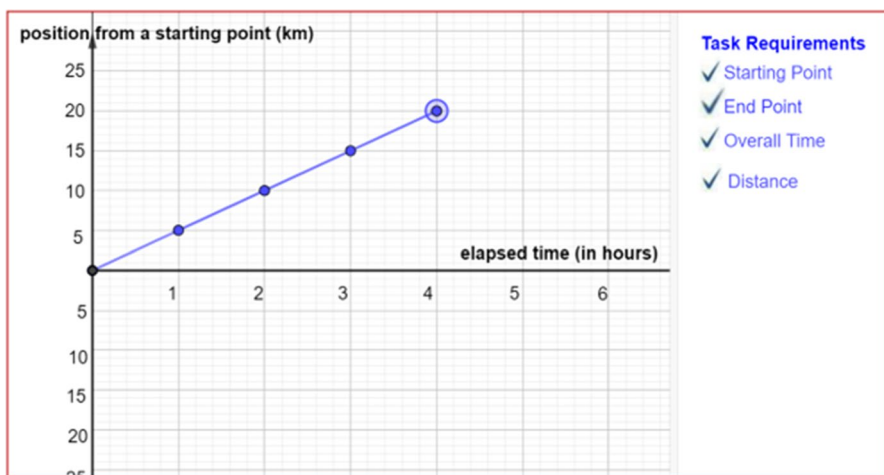

Fig. 6 A first example that is a Noga-trip

far [turn 9], going beyond the trials to produce the state of the diagram she did in Episode 1.

She formed a rule that she turned into a general procedure for constructing graphs that represent the situation, assuming that the starting and the end points should remain fixed points while others may change. Sarah approached this procedure by questioning her own intentions; she had probably thought that everything should change to create another example, then was convinced that the first point should be fixed at (0, 0) and then that the end point should also be treated as a fixed point.

8. Okay.
9. Well now it's really easy. I shouldn't touch the first [the left-most point]. It must be zero and zero of distance and time. In fact, why should I move the last point at all? All together [sums up the four segments], it has to take four hours and 20 km. So, I'll just move it a bit [moves the second point on the left] and this [moves the third point on the left and the fourth, and creates the middle drawing] … So. Okay.
10. Excellent. I'll save it.

Sarah creates the graph shown in the right-most chart (Fig. 4) in the same way.

11. I finished. I'm going to submit my answer and see what the system thinks of what I did

Sarah submitted two more states while observing the linked characteristics and their affirmation that she was submitting states of the interactive diagram that answer the requirements of the task. She had not yet received a post-submission report [turn 11] but had completed her work towards submission of the task.

Understanding the two functionalities of the mirror, as validating actions taken and guiding future actions, in this episode Sarah began to think about how to create multiple examples and respond to the requirement of the task for a variety of examples. She moved beyond the consideration of a single correct state of the interactive diagram and looked for a strategy to devise different examples. The strategy she came up with involved controlling one parameter and changing others (later in the vignette she is depicted as coming to realize the limitations of this strategy), which led her to generalize the actions she had taken so far [turn 9]. In doing that, Sarah envisioned distinctions among examples before constructing them. The mirror turned out to be a means for her to reflect on her as of yet unconstructed personal space of possible states of the interactive diagram.

At the same time, while the examples she created before receiving the post-submission report all met the requirements of the task, they left aspects of the example space unrepresented, and in that sense were too narrow. (Here, we are building on the use Lakatos, 1976, makes of the term "narrow" to describe the domain of a conjecture that is "safe," but leaves some examples outside of the scope of the conjecture.) The post-submission report provides a slightly different kind of mirror than the immediate feedback.

Episode 3: Reviewing the Characteristics in the Post-submission Report

So far, to submit three states of the graph Sarah had used the feedback within the interactive diagram to provide her with information. When she received the post-submission report (Fig. 4 presents a screenshot of Sarah's examples and the post-report she received) her first action was seeking confirmation that her submissions met the requirements, as she expected [turn 12]. And, she was anxious to see what else STEP said with the additional characteristics that were now available to her.

12. I see (in the post-submission report) that I managed to meet the requirements of the task but I have to look at those (characteristics)

Sarah searched through the list of the additional characteristics [turn 14]. As a result of having read through these, she seemed to have understood something that was new to her, perhaps concerning the possibility of changing direction or maybe even about the idea of the list of the additional characteristics [turn 17].

13. "Noga changed the direction of her riding." This is really interesting [scrolling down to view the rest of the list].
14. "Noga rode at different speeds." [reading a marked statement from the right column] Mmm … Why do they say that Noga rode at different speeds? No one said that we should construct at different speeds.
15. Maybe when the car stops it's considered zero speed?
16. "Noga changed direction at least once."

In answer to an interviewer's question about what she was doing Sarah said:

17. I'm just … looking at the gray [unmarked characteristics] I think about other things. … That's exactly it. That's what matters here in this list …

Having created three examples of Noga-trip, Sarah saved her work on task 1 and received a post-submission report. In contrast to the immediate information available, while she constructed her first Noga-trip, the description that Sarah received on her examples in the post-submission report was surprising. Rather than providing characteristics that determine whether or not Sarah's submissions were Noga-trips, she was now provided with other characteristics of the Noga-trips that she represented as graphs.

These additional characteristics brought forward new aspects of the situation that were not present in the task. For example, re-reading the task, Sarah emphasizes that, "No one said that we should construct a graph that shows Noga moving at different speeds." If the feedback provided by STEP is thought of as a mirror, it was a mirror that responded with a description that included characteristics that seemed out of place. For example, Sarah wondered why her examples caused the characteristic about different speeds to be highlighted.

Reflecting upon this feedback, she hypothesized that, "Maybe when the car stops it is considered zero speed [and therefore it is a change of speed]?" [turn 15] Indeed, two of her graphs showed a stop.

Sarah remained puzzled trying to understand what the mirror was now telling her. Leaving the change of speed aside for the moment she continued searching through the list and focusing on the challenging characteristic "Noga changed the direction at least once," which also was not identified anywhere in the description of the situation in the task, but that nonetheless was present in the feedback STEP provided her. Returning to the metaphor of a mirror that describes, the provided description highlighted something new and unanticipated.

Episode 4: Attending to Two Unexpected Characteristics

Next, Sarah worked on understanding what the characteristics "Noga finished riding in the city from which she left" and "Noga passed through the city" said about her submission.

18. So … it's another option. That she drove back and forth. This means that here [points at the right point] it's supposed to be on the *x*-axis because its distance is zero.
19. Ah … "Noga passed through the city" … so, she went through town … And probably she kept going negative.

Reading the report, Sarah acquired something that the mirror was telling her, and that she had not realized earlier: there are other options when riding, for example, riding away from the starting point and toward it. This idea of direction helped her settle her uncertainty about returning or passing through where the biker started.

As she worked with the post-submission report, she began to use the feedback from STEP as a tool to generate new ideas and understanding. The mirror that described her examples using unexpected words led her to re-analyze her assumptions and then to broaden her example space in several ways.

Episode 5: Describing Directional Noga-Trip Rides

Sarah now worked across the examples. The issue of changing directions appeared to Sarah as an important aspect of the story that she had not thought about before, and she asked the interviewer whether it would be relevant to the next task as well [turn 20]. She then retold the story of Noga's ride, summarizing for herself its meaning, which involves change of direction [turn 22]. Sarah thought about this exact characteristic already [turn 16], but was not aware of the centrality of what the talking mirror was saying then.

Sarah said to the interviewer:

20. Indeed I didn't consider it (changing the direction) as one of my examples.

21.   Does it mean that, in the next task, we want Noga to appear here as well [points under the *x*-axis]?
22.   But all that means that we'll have a change of direction [of the ride].
23.   I didn't construct any changes of direction [in my submitted example]. So, she [Noga] basically couldn't travel and come back.

Based on her uncertainty about why some of the characteristics were deemed important by the system, Sarah conducted another round of inquiry into the meaning of the distance traveled and of the end point [turns 24–27], which led her to conclude why her initial assumption about a fixed end point was limited [turn 27].

24.   … but if she goes back then her distance will no longer be 20 km.
25.   I don't understand…
26.   Then keeping the end-point as (4, 20) was a mistake? But that was the requirement, wasn't it?
27.   Ah. … Right. This is the total distance that she has to go through; the parts all together must be 20 km.

Sarah found that her conclusion about the direction change was a key aspect in the story that the system told her, and it had further implications; she needed to distinguish position from total distance traveled. She was now able to continue her story, telling that riding back and forth meant that Noga rode 20 km but did not necessarily reach the position of (4, 20). In this way, she completed the new story that she constructed listening to STEP which had listed the five characteristics and indicated their appearance or absence in her submitted Noga-trips. As with changing speed and stopping, she came to realize that Noga finishing where she began and passing through the city were both related to changing direction [turns 19–21]. These reflections on her part suggest that she may have achieved a central learning goal of the task, incorporating direction into her understanding of the situation.

Episode 6: Moving to the Second Task

At this point, Sarah felt ready to move on to the next task in the activity. She immediately acknowledged that the task was similar to the first one, but with fewer segments in the trip, and she quickly constructed three examples (see Fig. 7).
Sarah said to the interviewer while constructing these examples:

28.   So, I was trying to make three as different as possible. I started by changing the slope [pointing at the left-most example in Fig. 7] and then I made it to cross the *x*-axis … Yes, because it's about … She changed and went negative … She rode in one direction [pointing at the second example] 5 km, and then changed direction and then the distance is under here.

Sarah constructed the graphs to make it go "negative" below the *x*-axis as she had planned while working on the first task [turn 21].

Fig. 7 Sarah's graphs and post-submission report for task 2

Episode 7: Using the Additional Characteristics in the Post-submission Report to Track Variety in Her Example Space

Sarah examined the report and was happy to get confirmation that she created three different examples, as the task required.

29. In the post-report of the first task only two characteristics were highlighted. Now I added five highlighted characteristics.
30. But also look here, they vary in each example … almost different. Two here, in the first, are not the same two here (on the third), and the second on the row (stopping) appears only here and not in the other two. So altogether I diversified the examples a lot!

Her example space in response to task 2 was substantially broader than her submissions for task 1. She moved from three examples that all began and ended at the same places to three examples that began in the same place, traveled the same distance but, as a result of shifting directions, ended in different places. She also now seemed to view the additional characteristics as indicating differences among Noga-trips, describing what STEP told her through the number and diversity of the highlighted markings.

Supporting Students' Reflections on their Own Understandings

As presented in the previous section, the vignette captures how the students interviewed by the third author used the information they were given by STEP to create feedback processes that helped them learn. In terms of the feedback processes, as illustrated in the vignette, students explored what the pre-submission immediate feedback told them about the graphs they had created as possible Noga-trips, first looking at one example and then because of the requirement to submit multiple examples across examples. Thus, at first, the focus of the exploration was on what graphs represented about Noga-trips and what graphs did not. The requirement of submitting multiple examples began to give the interviewer a sense of the students' personal example space, but the students were not yet reflecting on the nature of their personal example spaces; they were focused on whether or not they had created Noga-trips.

The focus of their feedback process changed when they received the post-submission feedback, which both told them that they had created a set of Noga-trips and seemed to indicate that those trips did not exhaust all of the possible kinds of such trips. Students then worked on understanding the role of direction and of the endpoint of the trip in creating the possibility for other Noga-trips that were different from the ones that they had initially created. Moving on to the second task, as the vignette illustrates, students were able to generate Noga-trips quickly and to generate a more diverse personal example space than on task 1. We take this as evidence that there was a transition in their thinking about total distance traveled.

One way to consider what students may have learned from feedback provided in instructional environments is to examine whether the information provided as feedback seems helpful in reaching instruction goals and whether the information received as feedback seems to have the potential to improve students' performance on similar tasks. With these criteria in mind, we summarize our understanding of how the vignette we have presented answers the research questions and represents the ways that description in words helped students learn about the representation of a motion situation with graphs.

In the presented vignette, when working on the first task, the student submits a series of graphs in which the meaning of the feedback "Overall distance meets the task requirements" was that Noga ended her ride 20 km away from her starting point. There were no submitted examples in which she had changed direction, in other words, there were no examples where the total distance Noga traveled was 20 km and where she ended up less than 20 km from her starting point. After students submitted examples of Noga-trips, the post-submission feedback that included the possibility that "Noga changed direction at least once" caused the student to explore graphs involving a change of direction and led to the realization that in such trips, Noga could still have biked 20 km, even though she would end up less than 20 km from her starting point. When returning to the interactive diagram in task 2, the immediate feedback that nonetheless for such a graph the "Overall distance meets the task requirements" is a resource for a transition in

understanding that involves the development of nuance in the students' understanding of distance in the source path goal schema.

At first glance, the immediate information that students received is what a student might receive from online feedback designed in the image of Multiple Linked Representations: an immediate-linked description of the graph they submitted in words of the situation (distance, starting and end points, and time). This feedback helped students explore the situation, but also to understand and respond to the task requirements. In a conventional problem to solve, such descriptive information would be conceptualized as helping students come to and validate correct answers. But the task students were given requires that students construct several different examples, not just a single correct graph. By interactively trying out different positions for the draggable points on the graphs that students had ways to improve their understanding of the reference points of the source-path-goal schema: the starting and end points, and the representation of elapsed time and distance. According to the literature, this is not a trivial lesson for students to learn on their own while solving a motion task (Lakoff & Núñez, 2000; Radford, 2009).

Thus, in the vignette, we see the student use the feedback on critical characteristics of examples for different purposes from the feedback on non-critical characteristics provided in the post-submission report. The interviewed students used the immediately provided feedback on the critical characteristics of a Noga-trip to determine whether or not their graphs exemplified the story in the task. At the same time, having the post-submission feedback followed by a return to the interactive diagram in the second task allowed students to explore whether or not their personal example space was too narrowly defined and could benefit from including a greater variety of potential Noga-trips.

If the students we interviewed had instead used an online feedback system that just analyzed how their answers met requirements, they might have continued directly from task 1 to task 2. But, in the interviews, before students moved on to the next task, the STEP activity offered them a post-submission report with feedback information individualized to their specific examples. By contrast with the pre-submission immediately provided information, the information in the post-submission report was unexpected.

As captured by the vignette, working with the characteristics that helped them distinguish among Noga-trips, the interviewed students were able to reflect on their submissions to task 1, to realize that their personal example spaces could be expanded, to generate broader example spaces for task 2 and then to use the post-submission report for task 2 to validate the fact that the submitted examples capture a broader example space. Each example activated different additional, non-critical characteristics and, collectively, the three examples activated all the available characteristics, in that sense of representing a broader personal example space.

Among the major processes that were helpful in the learning supported by the post-submission report that is illustrated in the vignette, we identified learning to understand the difference between critical characteristics that distinguish between examples and non-examples and the non-critical characteristics that capture differences among Noga-trips. Beyond that, the vignette captures how understanding several separate characteristics together signals an important mathematics aspect of the situation in the task that students had not considered before the direction of travel,

which also has implications for the end-point. Thus, the vignette in this article offers initial evidence to suggest that supporting a student's inquiry with a mirror that speaks is feasible, and it identifies the feedback processes that we saw among our interviewees.

Reflecting on Ideas for Future Design Development

We close this article by reflecting on the important role that both critical and non-critical characteristics have played in our enactment of providing students descriptions of their examples in words and questioning design considerations of online feedback platforms that deserve further examination.

What we have found in this research study suggests that task designers can design tasks to help students examine their personal example spaces and shift their understanding, resulting in a broader personal example space. By adroitly using an analysis of tasks using Hershkowitz's categories of critical and non-critical characteristics, task designers can fashion tasks with the automatic characterization of students' examples in words. These characterizations provide information to students about their examples from which students may construct a feedback process that will include reflection on their personal example spaces.

More specifically, the vignette illustrates that what Hershkowitz (1990) calls *critical characteristics* (she focused on concepts) helps students track whether an example is an example of a concept or, as illustrated in this article, is responsive to the request for examples that a task makes (for a graph of a Noga-trip). These critical characteristics distinguish between examples and non-examples, while what Hershkowitz labels as non-critical characteristics can be used to distinguish among examples of a concept and thus can help with students' concept image or personal example space. This observation is an important one for designers of automatic feedback to use to achieve their pedagogical goals.

Of course, we can never collect the full variety of the examples in students' personal example space. The states of the interactive diagram that students submit in response to a task contain the first three examples that occur to students. These submissions do not include examples that do not come as quickly to mind and that students could possibly consider submitting if the task would allow for more examples. In addition, tasks of the kind we have looked at here do not explicitly ask for submissions of non-examples (by way of contrast, see the task in Yerushalmy et al., in press).

There are other design questions as well, for example, concerning when the feedback is given. In the tasks used in the interviews reported upon in this article, all of the critical characteristics were given in the immediate feedback and key additional characteristics were given only in the post-submission report. By studying these tasks, we are not claiming that this is the only way to proceed. Indeed, a recent article describes tasks in which some non-critical characteristics are available in the immediate feedback (Yerushalmy et al., in press). Our instinct suggests that pedagogical goals and location in a learning trajectory will determine such strategic choices, but this matter deserves careful study.

Similarly, unlike the procedure followed in the interviews, we report on here, in the way that STEP currently works, the platform does not require students to view or reflect on post-submission reports before going on to the next task in an activity. Indeed, students do not need to complete a particular task before working on the next one. We believe that, again, learning goals will determine the effectiveness of different strategic choices. For this reason, we continue to study how to design feedback information and processes to support student learning from one task to another, and also from one activity to another.

At the same time that we continue to refine the STEP platform, and our understanding of how to design tasks and the sorts of automatic feedback described in this article, as mathematics educators we fully acknowledge that the setting of an individual student facing a digital platform is just one piece of the learning puzzle and is a setting that is limited in terms of supporting even the learning process of an individual student. Focusing only on such settings overlooks important aspects of learning. Non-judgemental descriptive feedback can be used by teachers to support student inquiry processes. Efforts to use the descriptive feedback that we have outlined to promote self-reflection as part of the learning process need to explore other classroom settings, like working in pairs or small groups, and to incorporate teachers as facilitators of the learning process in and out of the classroom.

Data Availability The authors confirm that the data supporting the findings of this study are available within the article.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Bokhove, C., & Drijvers, P. (2012). Effects of a digital intervention on the development of algebraic expertise. *Computers & Education, 58*(1), 197–208.

Carless, D. (2015). *Excellence in University Assessment: Learning from award-winning practice*. Routledge.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325.

Chazan, D., & Ball, D. (1995). *Beyond exhortations not to tell [microform]: The teacher's role in discussion-intensive mathematics classes* (NCRTL Craft Paper 95–2). Michigan State University.

Gilead, S., & Yerushalmy, M. (2006). Graphs that are close to situations: Affordances and constraints. Talk presented at the 2006 *American Educational Research Association conference* (https://www.researchgate.net/publication/255658470). Accessed 13 Dec 2022.

Goldenberg, P., & Mason, J. (2008). Spreading light on and with example spaces. *Educational Studies in Mathematics, 69*(2), 183–194.

Harel, R., Olsher, S., & Yerushalmy, M. (2022). Personal elaborated feedback design in support of students' conjecturing processes, *Research in Mathematics Education*. https://doi.org/10.1080/14794802.2022.2137571.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hershkowitz, R. (1990). Psychological aspects of learning geometry. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and cognition* (pp. 70–95). Cambridge University Press.

Hohenwarter, M., Jarvis, D., & Lavicza, Z. (2009). Linking geometry, algebra, and mathematics teachers: GeoGebra software and the establishment of the International GeoGebra Institute. *International Journal for Technology in Mathematics Education, 16*(2), 83–87.

Jensen, L., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning: A critical review and metaphor analysis. *Computers & Education*, *173*, (104271).

Laborde, C., & Laborde, J.-M. (2014). Dynamic and tangible representations in mathematics education. In S. Rezat, M. Hattermann, & A. Peter-Koop (Eds.), *Transformation: A fundamental idea of mathematics education* (pp. 187–202). Springer.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge University Press.

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

Mason, J. (2008). PCK and beyond. In P. Sullivan & T. Wood (Eds.), *Knowledge and beliefs in mathematics teaching and teaching development*: *International handbook of mathematics teacher education* (vol. 1, pp. 301–322). Sense Publishers.

Naftaliev, E., & Yerushalmy, M. (2017). Engagement with interactive diagrams: The role played by resources and constraints. In A. Leung & A. Baccaglini-Frank (Eds.), *Digital technologies in designing mathematics education tasks: Potential and pitfalls* (pp. 153–173). Springer.

Nathan, S., & Koedinger, K. (2005). Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational Psychologist, 40*(4), 257–265.

Nemirovsky, R. (1994). On ways of symbolizing: The case of Laura and the velocity sign. *The Journal of Mathematical Behavior, 13*(4), 389–422.

Olsher, S., Yerushalmy, M., & Chazan, D. (2016). How might the use of technology in formative assessment support changes in mathematics teaching? *For the Learning of Mathematics, 36*(3), 11–18.

Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics, 70*(2), 111–126.

Schnepp, M., & Chazan, D. (2004). Incorporating experiences of motion into a calculus classroom. *Educational Studies in Mathematics, 57*(3). (https://terpconnect.umd.edu/~dchazan/SchneppVP/). Accessed 13 Dec 2022.

Schwartz, J. (1989). Intellectual mirrors: A step in the direction of making schools knowledge-making places. *Harvard Educational Review, 59*(1), 51–62.

Schwartz, J., & Yerushalmy, M. (1995). On the need for a bridging language for mathematical modeling. *For the Learning of Mathematics, 15*(2), 29–35.

Sfard, A. (2007). When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint. *The Journal of the Learning Sciences, 16*(4), 565–613.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.

Talmon, V., & Yerushalmy, M. (2004). Understanding dynamic behavior: Parent–child relations in dynamic geometry environments. *Educational Studies in Mathematics, 57*(1), 91–119.

van der Kleij, F., Feskens, R., & Eggen, T. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475–511.

Yerushalmy, M., Chazan, D. & Olsher, S. (in press). Automatic reports to support students with inquiry learning: Initial steps in the development of content specific learning analytics. In W. Jianpan (Ed.), *Proceedings of the 14th International Congress on Mathematical Education: Invited Lectures.* World Scientific Publishing House.

Yerushalmy, M., & Gilead, S. (1999). Structures of constant rate word problems: A functional approach analysis. *Educational Studies in Mathematics, 39*(1–3), 185–203.

Yerushalmy, M. (2020). Seeing the entire picture (STEP): An example-eliciting approach to online formative assessment. In B. Barzel, R. Bebernik, L. Göbel, M. Pohl, H. Ruchniewicz, F. Schacht & D. Thurm (Eds.), *Proceedings of the 14th International Conference on Technology in Mathematics Teaching – ICTMT 14* (pp. 26–37). (https://duepublico2.uni-due.de/receive/duepublico_mods_00070728). Accessed 13 Dec 2022.

Zaslavsky, O., & Zodik, I. (2014). Example-generation as indicator and catalyst of mathematical and pedagogical understandings. In Y. Li, E. Silver, & S. Li (Eds.), *Transforming mathematics instruction: Multiple approaches and practices* (pp. 525–546). Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.