Educ Stud Math (2007) 65:349–365
DOI 10.1007/s10649-006-9054-0

Triadic interaction in clinical task-based interviews with mathematics teachers

Boris Koichu · Guershon Harel

Published online: 2 February 2007

Abstract A clinical task-based interview can be seen as a situation where the interviewer–interviewee interaction on a task is regulated by a system of explicit and implicit norms, values, and rules. This paper describes how documenting and mapping triadic interaction among the interviewer, the interviewee, and the knowledge negotiated can be used to increase procedural replicability of the interview and accuracy of drawn conclusions about the interviewee's thinking process. Excerpts from interviews with 25 inservice mathematics teachers working on a task to make up a problem whose solution requires division of two fractions are discussed. The excerpts illustrate the relationship between methodological decisions taken by the interviewer during the interview and the applicability of the interview output to the research questions. A divergent analysis of the interviews with these teachers, which spanned over two years and were conducted by four interviewers, is used to offer a framework for analyzing data collected in clinical task-based interviews.

Key words clinical task-based interviews · division of fractions · experimental contract · grounded theory · interactions between interviewers and interviewees · procedural replicability · protocol analysis · thinking aloud · posing word problems

1 Introduction

Clinical task-based interviews, where a subject talks during or immediately after solving a problem, are used by many researchers in mathematics education. Since the mid-1970s, when Piaget pioneered the method, it has evolved into a variety of techniques, including open-ended prompting and structured think-aloud protocols (e.g., Clement, 2000). There is overwhelming evidence that clinical task-based interviews open a window into the subjects' knowledge, problem-solving behaviors, and reasoning (e.g., Newell & Simon, 1972; Schoenfeld, 1985, 2002). It is also well-known that methods of clinical interviewing bear

B. Koichu (✉) · G. Harel
Department of Mathematics, University of California, San Diego,
9500 Gilman Drive MC 0112, La Jolla, CA 92093, USA
e-mail: bkoichu@math.ucsd.edu

powerful limitations, some of which are unavoidable, whereas others can be reduced or bypassed when special methodological and analytical efforts are made (Goldin, 2000).

This paper is drawn from our experience of making such efforts during a study dealing with the development of the knowledge base of middle-school algebra teachers. We will refer to this NSF-funded study as the DNR study. The acronym DNR denotes a system of pedagogical principles – Duality, Necessity, and Repeated Reasoning – which constituted a theoretical framework for the study; the interested reader can find the discussion of this framework elsewhere (Harel, 2001, in press a, in press b). This paper offers an approach to increase reliability of the interview data and accuracy of the inferred conclusions about the subjects' thinking by close analysis of a *triadic interaction* — an interaction among interviewee, interviewer, and the knowledge under consideration.

In Section 2 we discuss selected issues in clinical interviewing. This is followed, in Section 3, by brief description of the DNR study. Examples of the interviews from the DNR study and their generative analysis are placed in Section 4. In Section 5 we present the exploratory Interview Communication Map (ICM) — a framework for analyzing data collected in the interviews. Concluding remarks suggest how the presented ideas can be extended and used to make a methodological side of mathematics education research more explicit.

2 Methodological issues in clinical interviewing

Three methodological issues associated with clinical task-based interviews, which are broadly discussed (yet not resolved) in the literature about clinical interviewing (e.g., Clement, 2000; Ericsson & Simon, 1993; Goldin, 2000), are particularly relevant to this paper:

(a) What are the effects of the interviewer's instructions and prompts on the interviewee's actions?
(b) What is the role of procedural replicability in clinical interviewing?
(c) How does an experimental contract between researchers and subjects affect conducting the interviews?

2.1 What are the effects of the interviewer's instructions and prompts on the interviewee's actions?

One interviewer's monologue, recommended by Ericsson and Simon (1993) for initiating a subject's thinking-aloud talk, includes the following:

> [T]ell me EVERYTHING you are thinking from the time you first see the question until you give an answer. I would like you to talk aloud CONSTANTLY from the time I present each problem until you have given your final answer to the question. I don't want you to try to plan out what to say or try to explain to me what are you saying. Just act as if you are alone in the room speaking to yourself. It is most important that you keep talking. If you are silent for any long period of time I will ask you to talk (p. 378).

Based on the analysis of many interview studies, Ericsson and Simon suggest that such instructing a subject to talk is likely to elicit a valid verbal report on his or her thought processes. Validity of a verbal report corresponds to the extent to which subject's talk

represents the actual sequence of thoughts mediating solving an interview task (Clement, 2000; Ericsson & Simon, 1993).

From the above monologue, one gets the impression that the interviewer anticipates difficulties with the subject following the instructions. Indeed, the negative assertion "I don't want you to try to plan..." and the warning "If you are silent..." suggest that deviations from the intended thinking-aloud procedures are expected. This raises a question: What should one do if the subject does not think aloud as instructed? At this point, the Ericsson and Simon's guideline is less prescriptive than it was regarding the introductory part of the interview. They recommend correcting possible deviations by simply repeating key phrases of the general instruction, and if needed, by interspersing silence with a reminder: "Keep talking."

Many authors point out that prompts aimed at encouraging the interviewee to think aloud or to clarify his or her actions can have unexpected and undesirable effects. For example, Ericsson and Simon (1993) discuss the impact of the prompt "What are you thinking about?" on the subject's talk (p. 83). They argue that such a question is likely to lead the subject to a self-observation talk or to an interviewer-oriented description as a response. Aronsson and Hundeide (2002) point out that when an interviewer repeats a question, it is often conceived by the subject as request to withdraw the previous answer and substitute it with a new one. These observations imply that, while prompts to think aloud as well as clarification questions can help keep the subject talking and invoke important information, they also can disturb an interviewee and deprive him or her of an independent problem-solving experience.

Ericsson and Simon (1993) and Van Someren, Barnard, and Sandberg (1994) call for purity in the implementation of thinking-aloud techniques. However, Clement (2000) points out that there are many trade-offs in one's decision of how rigorous or flexible the interviewer should be during the interview:

> [T]he need for completeness argues for doing more probing, whereas the need for minimizing interference argues for doing less. A compromise appropriate for some purposes is to do only nonintrusive probing early on... and more intrusive probing only later in an interview. How one resolves this trade-off also depends on the purpose of the interview (p. 571–572).

While reflecting in Section 4 on the interviews conducted during the DNR project, we distinguish between situations where any probe is undesirable as potentially distorting and other situations where silence on the part of the interviewer is less justified than intensive probing.

2.2 What is the role of procedural replicability in clinical interviewing?

Procedural replicability is a means of fostering the reliability of observations (Clement, 2000). It refers to the extent to which the procedure of presenting interview tasks and prompts to the subjects is specified explicitly enough to be replicable at different times and across different subjects and experimenters. The issue of procedural replicability is also related to generalizability of findings. As Clement (2000) notes, "Which procedures are standardized will depend on the issues that one wishes to reach conclusions about over groups of subjects" (p. 578). For instance, in the DNR project, we wish to reach conclusions about teachers' *ways of understanding* and *ways of thinking* while posing mathematical problems. Our concern about procedural replicability is determined, in part, by the

fact that a total of 25 teachers-participants were interviewed by four members of the research team, and interviews with 11 of 25 participants were repeated for two more times — after one and two years of the instruction.

It is particularly difficult to account for non-verbal events that (potentially) influence subjects' thinking during the interviews. Indeed, besides explicit instructions and prompts, many tacit rules regulate a dialogue between an interviewer and a subject. Schubauer-Leoni and Grossen (1993) observe that since the experimenters generally refrain from explicitly validating the subjects' answers, the subjects tend to base their auto-attribution of achievement on implicit clues like nods, smiles and conversational markers. Such indispensable attributes of any dialogue between human beings can hardly be standardized. Moreover, the harder the interviewer would try to wear the mask of neutrality, the harder the subject would try to figure out what is beyond the mask. This is because inevitable asymmetry in positions of the interlocutors in clinical task-based interviews: the interviewee assumes (and often rightly!) that the interviewer knows a solution to the interview task and thus conceives the interviewer not only as an addressee of responses but also as a possible source of assistance and evaluation.

Goldin (2000) points out that describing constraints or influences from the social contexts of clinical task-based interviews, as well as their individual psychological contexts as experienced by the subjects, is challenging and calls for investigation:

> Overt consideration and analysis of such issues, even when they are not the main focus of the research, can contribute substantially to the quality of the research study through improved interview design and more careful inferences drawn from the outcomes observed (p. 534).

The analysis of interviewer–interviewee interactions in Section 4 is definitely influenced by this call.

2.3 How does an experimental contract between researchers and subjects affect conducting the interviews?

A clinical task-based interview can be seen as a situation in which an interviewer and a subject interact on a task. Their interaction is regulated by a system of explicit and implicit norms, values, and rules — by an *experimental contract* (Schubauer-Leoni & Grossen, 1993). An assumption underlying the researchers' interest in different types of experimental contracts is that the relationship between the subject and the negotiated knowledge is mediated by social behaviors, which from the subject's perspective seem to be relevant to the situation. To interpret the nature of collected data, Schubauer-Leoni and Grossen (1993) recommend distinguishing between *private* and *public* constituents of the subjects' verbal responses. This recommendation influenced development of the Interview Communication Map, presented in Section 5 of this paper.

When describing adult–child interaction in a clinical interview, Aronsson and Hundeide (2002) point out that a subject may be seen attempting to please an interviewer in several different ways and to try "to be in the know", "to offer the 'correct' response", to "produce the interviewer's preferred response" (p. 181) etc. While Schubauer-Leoni and Grossen (1993) and Aronsson and Hundeide (2002) address mostly adult–child research settings, we suggest that their observations apply also to adult–adult interviews. Indeed, as is evident in Section 4, the interviews with mathematics teachers who took part in the DNR study are not free from actions that can naturally be attributed to the social component of the relationship

between the interviewers and the interviewees. We also show how the experimental contract is negotiated during a particular interview and how it varies from one interview to another.

3 The DNR study

For us, the need to deal in depth with the pros and cons of clinical interview methodology manifested itself in the framework of the DNR study. During two years of DNR-based instruction, teacher-participants were given an opportunity to be mathematical problem-solvers and to reflect on their problem-solving experiences from both students' and teachers' points of view. The main research question of this continuous and large-scale study is "What *knowledge base* do the mathematics teachers possess before, during, and after DNR-based instruction?" As defined in Harel (1993) and Harel (in press a), teacher's *knowledge base* includes:

- *Ways of understanding* mathematics content, including the actions of producing meaning/interpretation for a term, statement, or problem.
- *Ways of thinking* that govern ways of understanding, including problem-solving approaches and proof schemes.
- Knowledge of pedagogy, including an understanding of how to pose problems.

To address (partially) the above research question, the semi-structured clinical interview was designed. The interview task was to compose a word problem involving division of fractions. If and when a word problem was posed, the participants were prompted to discuss its mathematical and pedagogical aspects, and then to answer more general questions about division of fractions (see Fig. 1 for exact wording of the first three interview tasks).

The interview was used to detect components of the teachers' knowledge base in one particular context – fractions – and to address the following "local" research questions:

Before, after one year, and after two years of the DNR-based instruction,

1. What word problems involving division of fractions can the teacher-participants make up?

Task A

"Make up a word problem whose solution may be found by computing 4/5 divided by 2/3."

Task B

"Are you convinced that solution to your problem is 4/5 divided by 2/3? Why?"

"If you are convinced, how would you convince your students that the solution to your problem is 4/5 divided by 2/3?"

Task C

"How would (or did) you compute 4/5 divided by 2/3?"

"What makes you certain that the rule 'invert and multiply' works?"

"How would you explain to your students in a first year algebra class that the rule works?"

Fig. 1 The interview tasks

2. What ways of thinking seem to govern the teacher-participants' processes of constructing of word problems involving division of fractions?
3. What ways of understanding fractions and division do the teacher-participants possess?
4. How do the teacher-participants teach their students to make sense of division of fractions?

Using the interviews, we hoped to answer research questions 1 and 2 fully, and research questions 3 and 4 partially. It was clear to us that addressing the four "local" questions required attention both to ongoing reasoning processes of the teacher-participants and to relatively stable elements of their knowledge bases. Therefore, it was important for the thinking-aloud reports to be as undistorted as possible while the interviewees worked on Task A, and then to engage the participants in semi-structured conversations with the interviewer based on the questions of Task B and Task C. The interviewers were trained to follow closely the interview protocol, and, in particular, to refrain from revealing to the interviewees anything about the "correctness" of their responses. However, in spite of special efforts to standardize the key elements of the interview procedure, numerous between-interview differences due to different patterns of interactions among the interviewers, the interviewees and the knowledge negotiated during the interview were observed. That led us to formulate the following question, which we deal with in this paper: Given that the capability to standardize the interview procedure is limited, what analytical effort is needed in order to take into account the interview milieu in a way that reasonably assures accuracy of answers to the research questions of the DNR study?

4 Examples from the DNR study

We analyze here in some detail 30 minutes of excerpts from approximately 24 hours of audiotaped data collected in the interviews with 25 teacher-participants. The three interviewers are referred to as Int1, Int2, and Int3, and the subjects as Alon, Burt, and Carol. The chosen episodes are from four different interviews: the interviews with Alon and Carol are conducted by Int1 at the beginning of the project; one interview with Burt is conducted at the beginning of the project by Int2, and another is conducted after one year of DNR-based instruction by Int3.

The purpose of presenting these episodes is threefold: first, to show negotiating experimental contracts between the interviewers and the interviewees; second, to demonstrate trade-offs in conducting the interviews and between-interview differences; and third, to illustrate how analysis of the different types of interactions contributes to answering the research questions stated in Section 3.

Following Ericsson and Simon's recommendations for communicating verbal reports, we incorporate shortcuts and aggregations of the protocols whenever possible. However, relatively long excerpts from the transcripts are included in this section.

4.1 Negotiating of the experimental contract

After short mutual introduction, Int1, a research assistant in the DNR project, instructed Alon, the interviewee:

0:00 Int1: *We would like you to share your thinking with us. So we will ask you to think aloud about some*
0:30 *questions. Please understand that we may prompt you to continue thinking aloud at times. Don't worry about what we might think about your answers, we just want to learn about your thinking by seeing*

your honest responses to various questions. Now, the first question is, make up a word problem whose solution may be found by computing four-fifths divided by two-thirds [Int1 gives Alon a pen and a sheet of paper on the top of which Task A is printed].

This 30-second guideline opens the negotiation of the experimental contract between the interviewer and the interviewee. The guideline was designed to reduce a teacher-participant's stress at the beginning of the interview and to induce the subject to think aloud. As one can see, the interviewer's monologue essentially matches that quoted from Ericsson and Simon (1993, p. 378) in Section 2. The following dialogue takes place immediately after the above introduction.

0:32 Alon: *All right,* [Reading] *make up a word problem whose solution may be found by computing four-fifths divided by two-thirds. Um, let's see, and explain my thought process out loud as I go.*

0:49 *All right...Well the first thing I'd do is to find the solution. So divide four-fifths by two-thirds and so that* [Talking with pauses and writing] *is twelve-tenths which is sixth-fifths, which is one and one-fifth.*

1:23 *All right so, I am going to...*

1:23 Int1: *This is how you are solving, OK, OK.*

1:27 Alon: *Yeah...sorry would you like me to explain?*

1:30 Int1: *No, no, it's OK, just go ahead.*

1:32 Alon: *So I found the solution one and one-fifths um and let's see. If, and I'll just throw in these problems first.*

1:45 [Talking with pauses and writing] *Um, if Gloria picks a, why don't we say, four-fifths of all the apples*

3:25 *and divides them between herself and her two sisters, how much, er, how much, let's see... um, how many apples will Gloria and a sister have? OK, OK.*

3:33 Int1: *How would you convince your students that the solution to your problems is four-fifths divided by two-thirds?*

Consider the socio-cognitive events reflected in the above excerpt. Alon articulates the formulation of Task A and the key element of the instruction, says what he is going to do first ("...*to find the solution*") and does it. He is about to articulate his next step ("*All right so, I am going to*..."), but the interviewer interrupts him with the remark "*This is how you are solving, OK, OK*." Alon's next step remains unarticulated, since the interviewer's remark seems to distort the course of his thoughts. We think that this remark had different meanings for the two interlocutors. Several seconds before (0:49) Int1 was intrigued by the expression "*find the solution*" that Alon used when given the instruction to make up a word problem. When Int1 understands what Alon means by this, he reveals his delight by the above remark. On the other hand, Alon understands what Int1 has just said as a request to provide more detailed explanations. Int1 says "*No, no, it's OK, just go ahead*", but we observe that Alon changes his mode of speech from talking without signs of attention to the interviewer to addressing the interviewer, or, in terms of Schubauer-Leoni and Grossen (1993), from more private to more public talk. From this moment (1:32) Alon talks and writes with some attention to the interviewer and eventually produces what we will refer to as the Apple Problem (1:45). The assertion "*OK, OK*" (3:25) sounds like the end of the response to Task A, and Int1 turns to the second question of Task B (3:33) without any comments about the Apple Problem. As one can see, both interlocutors get used to their social roles and learn what to do and what not to do in the experimental format. As it is evident from the forthcoming fragment, implicit negotiation of the experimental contract between Int1 and Alon entails further adjustment of their behaviors and expectations.

During a time span of more than three minutes (3:34–7:01) Alon addresses Task B and explains to the interviewer how he would convince his students that the solution to the

Apple Problem involves dividing fractions; meanwhile, the interviewer remains silent. Suddenly, Alon stops himself:

7:02 Alon: *Wait, wait I'm not even convinced... Well it's kind of the, the pressure of the moment, you know. I'm kind of anxious about all this right now.*
Int1: *It's OK, haha... You know, it's just your spontaneous thinking, whatever that just comes to your mind.*
Alon: *Yeah, I'm comfortable being wrong...*
7:41 Int1: *So what is it that you are not convinced about?*
Alon: *I'm not convinced that my problem is correct...Because I haven't had a chance to check my work and, you know, you didn't tell me to make sure it worked.*
Int1: *You don't have to. It's up to you.*
7:57 Alon: *OK, so, I didn't really feel that it was absolutely necessary to check my work... I'd like to convince myself by working the problem out.*

Answering Task B, Alon figures out that the Apple Problem does not work. He explains that his mistake is due to the situation, and reveals his current understanding of the experimental contract, namely, the interviewee talks and the interviewer monitors. The phrase "*... you didn't tell me to make sure it worked*" reflects Alon's expectations that the interviewer would guide his actions and assure the success, even though he is "*... comfortable being wrong*". The interviewer's reply "*...It's up to you*" enriches the experimental contract by the new component: in spite of the unusual situation, the interviewee is fully responsible for results of his problem-solving.

After this conversation, Alon talks with no sign of attention to the interviewer, and the interviewer remains silent (7:57–12:35). Alon spends about 2 minutes trying to modify the Apple Problem, gives up, and then considers a new approach. Eventually, he produces, both orally and in writing, the following formulation, which we will refer to as the Money Problem: "*If sales* = 4/5 *total profit and cost is 2/3 of sales, what is cost?*" In writing, he connects this "word problem" to the expression $\frac{\frac{4}{5}x}{\frac{2}{3}}$ and says:

12:35 Alon: *I'm not really sure about this one either. I think I've kind of worked myself up to an anxiety point.*
Int1: *That's OK, OK. I don't mean 'good', I mean...*
Alon: *That I'm telling you where I am.*
12:42 Int1: *Yeah, exactly.*

From the last part of the interview (7:57–12:42), we learn that the two interlocutors have reached the desirable mutual understanding of the experimental contract. Namely, the interviewee not only cooperates and thinks aloud, but also does his best trying to solve the task without expecting the interviewer to guide him, and the interviewer does not deprive the interviewee of opportunities for independent problem-solving. Such a contract is likely to foster the reliability of the data — with respect to the purpose of the interview to investigate the subjects' *ways of understanding* and *ways of thinking*.

Now we address the natural question: How does the above analysis contribute to answering the research questions of the DNR study stated in Section 3? To us, the answer is twofold. First, we deem the above analysis useful for training the interviewers engaged in the project. Second, without considering the interviewer–interviewee interaction in depth, we might miss the process of developing the experimental contract and wrongly estimate the reliability of different parts of the interview. In turn, we might misinterpret the participant's ways of understanding and ways of thinking.

Consider an example — a piece of content analysis dealing with the particular question: based on the above interview, which ways of understanding fractions does Alon possess (see research question 3 in Section 3)? Alon attempts Task A two times: the first attempt results in the Apple Problem and the second attempt results in the Money Problem. From the first attempt, we infer that Alon understands the fractions 4/5 and 2/3 as processes of equal partitioning rather than objects or measures of quantities; he also seems to understand fractions in terms of *part–whole* (i.e., m/n means *m* out of *n* objects). However, these findings are incomplete since Alon's thinking may have been distorted and since the desirable experimental contract has not yet been established. As we have shown above, at the beginning of the interview Alon seemed to be more concerned with the obligation to think aloud than with doing his best to attack Task A. Alon attempts Task A for the second time after answering Task B and engages in an unstructured conversation with the interviewer. However, we deem that this part of the interview is reliable, since the teacher-participant's talk is private rather than public at that stage and since the developed experimental contract obliges Alon to do his best without relying on the interviewer. The second approach enables us to gain additional information about the participant's ways of understanding fractions. Indeed, when making up the Money Problem, Alon constructs the numbers 4/5 and 2/3 as measures of the quantities "profit" and "sales", and tries to utilize these numbers as arguments of an arithmetic operation.[1] Hence, the way of understanding fractions as objects rather than processes is indicated. For the reasons discussed above, we deem this piece of the data a product of the participant's undistorted thinking on the interview task.

In the DNR project, evolution of the experimental contract is also observed when comparing different interviews involving the same subject. Consider the corresponding fragments from two interviews with Burt.

Int2, one of the authors of this paper, interviewed Burt at the beginning of the project. After receiving the initial instruction (the same instruction as that given to Alon and quoted above), Burt started thinking aloud on Task A and after a while asked Int2:

Burt: *I'm rambling on here, am I getting anywhere near?*

Int2: *That's the whole goal is for me to hear it... Just feel free to ignore me, all right. Ignore me completely. Think whatever thinking quietly, think out loud. But you did a good job so far in sharing with me your thoughts*.

After the interviewer's response, which is in spirit of Ericsson and Simon's (1993) recommendations (see Section 2.1), Burt resumed working on Task A. A year later, Int3, the other author of this paper, interviewed Burt using the same interview task. Burt remembered that a year ago he could not make up a satisfactory word problem involving division of fractions, but did not remember which approaches had he tried. After receiving the initial instruction, Burt started thinking aloud on Task A and after a while asked Int3:

Burt: *Do you understand what I am trying to describe?*

In response, Int3 kept silence, and Burt resumed working on Task A and thinking aloud. We point out the between-interview difference in Burt's interpretation of the experimental contract: in the first interview, Burt was expecting the interviewer to guide his actions, and in the second interview he paid attention to the interviewer as a "neutral" observer.

[1] The Money Problem is unsolvable and really involves multiplication, not division. Indeed, it can be interpreted as follows: $\text{sale} = \frac{4}{5} \cdot \text{profit}, \text{cost} = \frac{2}{3} \cdot \text{sales}$, and then $\text{cost} = \frac{2}{3} \cdot \frac{4}{5} \cdot \text{profit}$. Regardless of the "correctness" of the Money Problem, the process of its formulating reflects the teacher-participant's way of understanding fractions. More detailed analysis of the teacher-participants' ways of solving Task A is a subject of an additional publication.

4.2 The interviewer's methodological decisions and trade-offs

In this subsection we discuss in some detail the actual scenario of one difficult interview and consider other scenarios that could have resulted had the interviewer made other methodological decisions.

After receiving the initial instruction (the same instruction as that given to Alon and quoted in Section 4.1), Carol, the interviewee, does not start talking; she works on Task A in silence. The excerpts below represent the interview from the moment when Int1 breaks about 90 seconds of silence.

1:58 Int1: *What are you thinking?*

2:00 Carol: [Giggles] *My mind is going blank, haha. I was thinking something about going to like a party and you have a certain amount of people and then they have to divide up like a pizza or they divide up a pie. I'm still trying to think of how many people...how they would want to divide it up.*
[Silence]

2:10 Carol: *Um, I guess you could say that if um, you're at a party with five friends and only four people,*
2:32 *four out of the five people want to have pizza, but there's only um...but then you find out that somebody*
2:58 *already ate one-third of the pizza, how would, how many pieces would you each have?*

As one can see, the interviewer allows Carol to keep silence for a relatively long period of time, and then asks her "*What are you thinking?*" The reply "*My mind is going blank...*" indicates that Carol feels uncomfortable when telling her thoughts as they come. Instead, she answers the interviewer's question retrospectively: Carol does not explain what she *is* thinking but what she *had been* thinking so far. This episode is an important one for understanding the actual experimental contract between Int1 and Carol: the interviewee works on the interview task in silence and then explains to the interviewer what she has done. Perhaps, if Int1 would remind Carol to talk after 15–20 seconds of silence, she would start thinking out loud, but we doubt this possibility. Note that Carol giggled when answering the interviewer's question — she was evidently confused by the difficulty of the interview task (This point will be further supported in the next excerpt). Based on the interviewer's impression from the teacher-participant's nonverbal behaviors along with the full transcript, we think that Carol intended to share mainly her success in doing Task A, not the struggle.

After an additional period of silence (2:10–2:32), Carol explains to the interviewer what she has done so far (2:32–2:58). The pure – in terms of Ericsson and Simon's methodological recommendations – thinking-aloud interview might have ended at that point. Indeed, the interviewee failed to think aloud and instead came up with an answer to the given task and explained how it was developed. In this case, such limited evidence would leave us, the researchers, with a few tentative suggestions about Carol's thinking and a long list of unaddressed questions. For example, we could say that Carol utilized the context of sharing pizza, apparently employing her previous experience in conceptualizing fractions, but we could not say how she plans and monitors the process of constructing a word problem.[2]

[2] In the framework of the DNR project, the teacher-participants' ways of thinking that characterize a mental act of problem-posing are explored (see research question 2 in Section 3). In particular, we are interested in reference points — pieces of knowledge that the teacher-participants hold as true and use as anchors in a processes of planning and monitoring of constructing word problem involving division fractions. We mention here two such reference points: ***Given task as a reference point***: This reference point is indicated when the teacher participant talks about the extent to which his or her formulation satisfies conditions of the interview task — without prompting from the interviewer. ***The answer as a reference point***: This reference point is indicated when the teacher-participant uses the observation that the result of dividing 4/5 by 2/3 is greater than 4/5 for evaluating whether or not a particular context is suitable for making up a problem involving division of the given fractions.

Indeed, it looks like Carol planned her formulation in silence and did not have a chance to monitor it, as was detected at the beginning of the interview with Alon. Fortunately, Int3 does not stop this part of the interview.

3:01 Int1: *Do you want to write that down?*
Carol: *Write that down. OK. Now I have to remember, haha.*
[Writing in silence: *You go to a party with friends. 4 out of 5 of you want pizza. When you open the box you find that 1/3 of the pizza is gone.*]
When I write it down it doesn't work, haha.
Int1: *What does not seem right?*
4:20 Carol: *Well, you have four out of five people but... it's not going to be four-fifths divided by two-thirds to get the solution...I'm thinking.*
[Silence]
Int1: *What have you tried so far?*
4:25 Carol: *I feel like I should be able to do this without having trouble.*
4:52 Int1: *Please, don't be stressed out about it... It's OK; just share with me... what you have been thinking ever since you felt that the solution to this* [Int1 points to the above written formulation] *is not four-fifths divided by two-thirds...*
6:04 Carol: *Um...Like, I was trying to think of somehow saying four-fifths of the people, or something like*
6:06 *that. Of four-fifths of the...like a total number.* [Talking with pauses and writing] So *you...maybe you*
6:25 *could say 4/5 of, um, 20 people at a party want to eat pizza but only 2/3 of them want to eat the pizza*
6:50 *now. How many are going to eat now? This might work better.*

In the above episode, Carol writes her first formulation (3:01) and, while writing, reconsiders and rejects it. After a while she produces a new formulation, which we will refer to as the 20-People Problem (6:25). From this segment of the interview we learn that Carol monitors her actions by referring to the given formulation of the interview task (see Endnote 3) and interprets 4/5 as a measure of the quantity "20 people". However, let us remember that this piece of data emerged from the interviewer's "technical" requests to write the previously formulated word problem. For this reason, we cannot confidently decide to what extent Carol's actions reflect her undistorted ways of understanding and ways of thinking. Perhaps she reconsidered her first formulation just because she was given the additional time, but it is also possible that she did so since she interpreted the interviewer's request to write as a tacit recommendation to change her mind (see discussion of Aronsson & Hundeide, 2002, in Section 2.1). Note also the assertion "*I feel like I should be able to do this without having trouble*." To us, it highlights Carol's confusion and supports the above point that Carol feels uncomfortable when talking about uncompleted or unchecked word problems. In summary, given the halting talk, we can hardly imagine what Int1 could have done differently in order to improve the actual situation: indeed, he just asked Carol to write her problem down in order to compensate for the poor verbal report with complementary evidence.

From the next four minutes of the transcript we learn that Carol is not sure that the 20-People Problem involves division 4/5 by 2/3 but can hardly capture what might be wrong in this "problem". During this time, she remains silent for relatively long periods of time (15–60 s), does not write, and incoherently responds to the interviewer's reminders "*Let me know what you are thinking*". At some point (11:20), the interviewer decides to change the course of the interview.

10:42 Carol: *It's not going to be exactly four-fifths divided by two-thirds to get the answer.*
Int1: *Well, why not?*
10:45 Carol: *Because twenty is in there.* [Rereads the formulation of the interview task]...*whose solution may be found by computing 4/5 divided by 2/3... First you'd have to take four-fifths of the twenty*

people and then two-thirds of whatever that number is...
[Silence]
Int1: *So if there were no "twenty"* [in the 20-People Problem], then it would work?
11:05 Carol: *Right, how many four-fifths of people.... Well you can't just say four-fifths of people; you have to know how many people...*
11:20 [Silence]
Carol: *I feel stuck, haha...*
11:26 Int1: *Oh that's OK. So how about if I would change...four-fifths of a GROUP of people. And then how...What FRACTION of the group of people? If I were changing it then would it work?*
11:46 Carol: *Yeah, then it would work.*
12:32 [Silence, Carol changes the written 20-People Problem and obtains the new formulation, which we will refer to as the Group-of-People Problem]: *"4/5 of a group of people of a party want to eat pizza but only 2/3 of them want to eat the pizza now. What fraction of the people will eat now?"*
12:44 Carol: *So four... What fraction...Yeah then I wouldn't...Maybe explain that if...You have to take four-fifths of the people and you'd have to divide it by two-thirds of them, because two-thirds of them don't want to eat it in order to plan your answer which is how many do want to eat it now...*
12:55 [Silence]
Carol: *Hmp*.
Int1: *Why? Why did you say 'hmp'?*
13:10 Carol: *Well six-fifths of the people will only eat pizza now, if I did that correctly, haha.*
13:44 Int1: *Is something bothering you?*
15:23 Carol: *Well I was thinking six-fifths is one and a one-fifth of the people. Doesn't really make sense.*
15:38 Int1: *What is it that doesn't make sense?*
15:48 Carol: *One and one-fifth of the people...That's like more than all of them, haha.*

In the above episode, we point to the major trade-off of the interview procedure: based on the information about the participant's difficulty with the 20-People Problem (10:45), the interviewer HELPS her to formulate the Group-of-People Problem (12:44–13:10). We do not accept the latter problem as a part of the answer to the first research question (see Section 3), since Carol did not produce it independently. However, the above episode contributes to the second and the third research questions. For instance, Carol spontaneously replies to the interviewer's suggestion: "*Well you can't just say four-fifths of people...*" (11:26). From this assertion we learn that so far Carol makes sense of the fraction 4/5 as a multiplier applicable only to a specified quantity. Apparently, this way of understanding fractions is a (partial) answer to the natural concern of why the interview task appeared to be so difficult for Carol. Moreover, her undistorted thinking aloud about the Group-of-People Problem (13:10–15:48) enables us to enrich the answer to the second research question by indicating an additional way of thinking. Namely, Carol uses the observation that the result of dividing 4/5 by 2/3 is greater than 4/5 for monitoring the Group-of-People Problem (see Endnote 3).

In closing the methodological discussion of the interview with Carol, let us note that the interviewer injected intrusive prompts when there was no hope that the subject would think aloud as planned, the trade-off being that while some pieces of the elicited information are less reliable than others, the collected data are rich with respect to the purpose of the interview: to learn about Carol's ways of understanding and ways of thinking in a particular mathematically-loaded situation. If we were given the opportunity to re-interview Carol, we could try to establish a better experimental contract with her. However, in our opinion, most of the interviewer's methodological decisions and trade-offs during the interview with Carol are vindicated.

5 Interview communication map

The Interview Communication Map (ICM) is an exploratory frame for taking into account triadic interactions in clinical task-based interviews. The ICM summarizes our attempts to find patterns in socio-cognitive events that took place in 47 interviews conducted in the DNR project, including those presented in Section 4. The ICM consists of *dimensions* and *categories.* According to Dey (1999, p. 252) these are analytical entities of a grounded theory: *dimension* is used to measure extensions; *category* is used as a way of identifying or distinguishing something based on comparison with other things.

Three *dimensions* of the ICM are *Subject*, *Interviewer*, and *Knowledge negotiated*. The first and second dimensions comprise six categories each and the third dimension comprises four categories. The names of the categories and their definitions are presented in Table 1.

The first and the second dimensions deal with a dialogical constituent of the interview. Each subject-oriented category distinguishes a certain level of observed attention that the subject pays to the interviewer. The interviewer-oriented categories describe the actions of an interviewer as an interlocutor who is in charge of initiating knowledge negotiation. These two dimensions of the ICM are interrelated yet asymmetric. The third dimension of the ICM deals with *potential* relevance of the participant's responses to answering the research questions, whereas *actual* relevance of the responses is decided by the analysts based on the three dimensions taken together. As one can see, assertions potentially useful as indicators of the subject's ways of understanding and ways of thinking are not separated. This is due to the Duality Principle, which is a part of the DNR conceptual framework (Harel, 2001):

> *Students' ways of thinking impact their ways of understanding mathematical concepts. Conversely, how students come to understand mathematical content influences their ways of thinking.* (p. 207; see also Harel, in press b)

In accordance with the ICM notation, three codes are assigned to each interview fragment. For example, the first five minutes of the interviews with Alon and Carol are coded in Table 2.

One needs some practice in order to start reading the ICM applied to particular interviews, but with practice, tables like Table 2 become useful representations of the interview data. In part, the ICMs can be unpacked. For example, one can learn from Table 2 that about 1:30 after beginning the interview Alon started to formulate a word problem and that his talk included potentially useful information about his ways of understanding and ways of thinking. However, at that point Alon started to pay attention to the interviewer. Since before 1:30 he talked without paying attention to the interviewer, we conclude (even without looking at the transcript in Section 2.3) that the interviewer's intervention was not necessary. We then see that Alon changed his mode of speech when answering the next interview task. Analogously, Carol's difficulties with thinking aloud and the corresponding interviewer's actions, discussed in Section 4.2, are abbreviated in the second part of Table 2. The table also highlights differences between these two interviews.

The presented ICM is context-dependent, but we believe that its developmental principles, elements, and structure can be adapted to various interview designs and research questions. At this point, we note that the interlocutors' actions are categorized in the ICM using not only the transcript of the interviews, but also authentic evidence such as

Table 1 Dimensions and categories of the ICM

Category	Definition
1. Subject (S)	
S1. Remaining silent	The subject remains silent for more than 15 seconds[3]. He or she seems to be engaged in doing an interview task.
S2. Soliloquy with no observed signs of attention to the interviewer	The subject talks for more than 15 seconds with no observed signs of attention to the interviewer. The talk is characterized by relatively low voice, uncompleted sentences, interjections (e.g., *um, hm, oh*) etc.
S3. Monologue with attention to the interviewer	The subject talks for more than 15 seconds, and the talk: (1) directly addresses the interviewer as an interlocutor (e.g., "*um, you know that...*", "*you asked me...*"), or (2) the talk is unusually (for the subject) deliberated and accompanied by nonverbal signs of attention to the interviewer of all kinds (e.g., intonation, looks etc.)
S4. Questions to the interviewer as a "neutral" observer	The subject breaks his or her action S1–S3 with requests: (1) to explain elements of the interview procedure (e.g., "*Can I write?*"), or (2) to provide feedback on clarity of actions (e.g., "*Do you understand what I am doing?*")
S5. Questions to the interviewer as a possible source of assistance or evaluation	The subject breaks his or her action S1–S3 with (1) requests for feedback on solution steps (e.g., "*Is it correct?*", "*Am I getting anywhere near?*"), or (2) asking for help (e.g., "*Can you help me?*")
S6. Responding to the interviewer's actions I4–I6	The subject responds to the interviewer's actions I4–I6. The responses are relatively short (as a rule, less than 15 sec.) and constitute the subject's part in dialogues initiated by the interviewer.
2. Interviewer (I)	
I1. Explaining the interview procedure and giving Task A	The interviewer: (1) explains the interview procedure at the beginning of the interview and gives Task A, or (2) responses to S4 (e.g., "*You can write*")
I2. Remaining silent	The interviewer observes the subject's actions in silence.
I3. Prompting to think aloud	The interviewer interrupts actions S1–S3 with short remarks (e.g., "*OK*", "*Go ahead*", "*Keep talking*" etc.) or interjections (e.g., "*mm*", "*aha*", "*oh*") aimed at encouraging the participant to think aloud.
I4. Requesting clarification	The interviewer interrupts actions S1–S3 with questions and remarks aimed at clarification of what an interviewee is doing, saying or thinking (e.g., "*I want to understand...say it again or write it down,*" "*What did you do?*", "*What are you thinking about?*", "*You divided it into thirds, right?*")
I5. Providing help or evaluation	The interviewer responds to S5 or interrupts actions S1–S3 with assertions aimed at: (1) providing problem solving clues (e.g., "*Try to divide it into thirds*") or (2) evaluating what the subject has done (e.g., "*It is correct!*")
I6. Giving additional tasks	The interviewer redirects interview by asking questions of Task B, Task C or additional interview tasks (see Fig. 1)
3. Knowledge negotiated (K)	
K1. Word problems	The subject's assertion that the analysts consider potentially relevant to the first research question (see Fig. 1).
K2. Ways of understanding and ways of thinking	The subject's assertion that the analysts consider potentially relevant to the second and third research questions.

Table 1 (continued)

Category	Definition
K3. Teaching fractions	The subject's assertion that the analysts consider potentially relevant to the fourth research question.
K4. Alien information	The subject's assertion that the analysts do not consider potentially relevant to the research questions.

[3] Consideration of 15-second intervals in the ICM is based on the Ericsson and Simon's (1993, p.83) remark: in different task-based interview studies the researchers reminded subjects to speak after 15 seconds to 1 minute of work in silence.

audibility of the subject's speech and intonations. We suggest that using a video camera trained on the interviewee is useful if the researcher plans to utilize the ICM.

6 Summary and concluding remarks

Clinical interviewing is an important instrument in current mathematics education research (Clement, 2000; Goldin, 2000; Schoenfeld, 2002). When properly implemented, it allows the researchers to open a window into the hidden world of the ways of understanding and ways of thinking of people doing mathematics. Even when properly implemented, though, the method is burdened by powerful limitations and constraints. To us, the need to deal in depth with some of the methodological constraints was apparent in the DNR study, in which the research question about development of mathematics teachers' knowledge base through stages of the DNR-based instructions was posed. The "local" research questions of the DNR study concern the teacher-participants' ways of thinking that govern the processes of constructing word problems involving division of fractions, their ways of understanding fractions, and their knowledge of how to teach division of fractions. Clinical task-based interviews with the teacher-participants were implemented to (partially) address these questions. In designing and administering the interviews, we followed recommendations by Clement (2000), Goldin (2000), and Ericsson and Simon (1993).

The cornerstones of the adapted methodology are to engage a subject in thinking aloud while solving a problem, to refrain from intrusive probing early on in the interview, and to involve the subject into semi-structured conversations with the experimenter later on in the interview. The methodology worked as expected. On the one hand, it supplied us with rich

Table 2 Alon and Carol: the fragments of ICMs

Time:	0:00	0:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00
Alon											
S:	S1	;S2		;S6; S3				;S2+S3			
I:	I1	;I2		;I4+I3; I2				;I6, I2			
K:	K4	;K2		;K1+K2				;K3			
Carol											
S:	S1				;S6; S1	;S3	;S1			;S6	;S1
I:	I1	;I2			;I3; I2		;I4 ;I2			;I4	
K:	K4				;K2	;K1+K2	;K1				

information about the participants' problem solving; on the other hand, it entailed many issues that are pointed out but not resolved in the literature. Specifically, because of the complexity of the matter investigated, conclusions about the teacher-participants' ways of understanding and ways of thinking in solving the interview tasks appear to be sensitive to unavoidable deviations from the "ideal" interview procedure. As Goldin (2000) suggested, each interview appears to be unique with respect to the interactions among a subject, an interviewer, and a problem, but this does not mean that researchers should give up and stop trying to filter the biases and to seek for quality interview data. Generally speaking, quality of the interview data depends on the nature of social and cognitive activities of the interlocutors, but the devil is in the details. The details of the experimental contracts regulating the course of the interview as well as of the circumstances of making particular methodological decisions by the interviewer are responsible for the outcome of the interview.

We demonstrated that the instruction articulated by the interviewer at the beginning of each interview constitutes only a part of the experimental contract. The actual experimental contract is gradually built by the interlocutors throughout the interview and from one interview to another. We demonstrated that the interviewer's interventions and trade-offs can have different consequences. Some of them can corrupt the reliability of the data, others can straighten the course of the interview, and there are additional probes that can have a mixed effect: they corrupt the data with respect to one research question and are beneficial with respect to the others. We also demonstrated how accounting for different types of social interactions between interviewee and interviewer can strengthen analysis of mathematical and cognitive content of the interview.

Furthermore, any research design attempts to reduce the complexity of the object under investigation (Schubauer-Leoni & Grossen, 1993). However, given that many variables influence the interview data, there is a risk of social or cognitive reductionisms in the data analysis (Boden & Zimmerman, 1991; Goldin, 2000). As a remedy, Schubauer-Leoni and Grossen (1993) suggest considering the triadic interaction among subjects, experimenters, and the knowledge in question to be an irreducible unit of analysis, whose functioning can only be understood within contextual dimensions. This idea is reified in our study as the Interview Communication Map. It appeared as a natural answer to the main question of this paper: Given that one's capability to standardize the interview procedure is limited, what analytical effort is needed in order to account for the interview milieu in a way that reasonably assures accuracy of answers to the research questions concerning development of subjects' ways of understanding and ways of thinking in a particular mathematical context?

To us, the ICM has four functions. First, it serves to shape the analysts' thinking about the nature of the interview data in designing and analyzing the interviews. Second, it provides useful notation, which helps to reduce the data and to focus on the most reliable and relevant (to the research questions) parts of the transcripts. Third, it demonstrates differences between the interviews. Fourth, it helps us to be better interviewers or at least to be more explicit regarding the quality of the data, since it highlights trade-offs and their consequences. Let us note that the ICM does not free the analysts from ad-hoc decisions regarding applicability of the data to the research questions, but it makes making the decisions more structured and explicit. This is especially important when analyzing large data sets, as in the DNR project.

In concluding, we suggest further exploring of the ICM-techniques. It seems to us particularly interesting and important to investigate how the ICM can be used for

aggregation and qualitative comparison of clinical interviews dealing with mathematical reasoning and mathematical problem-solving, and, probably, with other types of reasoning and problems.

Acknowledgments This work is supported, in part, by the National Science Foundation (Project # REC 0310128, G. Harel PI). Opinions expressed here are those of the authors and are not necessarily those of the Foundation. We also thank the anonymous reviewers for valuable comments and suggestions that influenced our thinking of the topic of this paper.

References

Aronsson, K., & Hundeide, K. (2002). Relational rationality and children's interview responses. *Human Development, 45*, 174–186.

Boden, D., & Zimmerman, D. (1991). *Talk and social structures*. Cambridge: Polity.

Clement, J. (2000). Analysis of clinical interviews: Foundation and model viability. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). New Jersey: Lawrence Erlbaum.

Dey, I. (1999). *Grounding grounded theory: Guidelines for qualitative inquiry*. San Diego: Academic.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data (Revised version)*. Cambridge: MIT Press.

Goldin, G. (2000). A scientific perspective on structures, task-based interviews in mathematics education research. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 517–545). New Jersey: Lawrence Erlbaum.

Harel, G. (1993). On teacher education programs in mathematics. *International Journal for Mathematics Education in Science and Technology, 25*, 113–119.

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zaskis (Eds.), *Learning and teaching number theory* (pp. 185–212). New Jersey: Ablex.

Harel, G. (in press a). What is mathematics? A pedagogical answer to a philosophical question. In R. B. Gold & R. Simons (Eds.), *Current issues in the philosophy of mathematics from the perspective of mathematicians*. Mathematical Association of America.

Harel, G. (in press b). The DNR system as a conceptual framework for curriculum development and instruction. In R. Lesh, J. Kaput, E. Hamilton, & J. Zawojewski (Eds.), *Foundations for the future*. New Jersey: Lawrence Erlbaum.

Newell, A., & Simon, H. (1972). *Human problem solving*. New Jersey: Englewood Cliffs.

Schoenfeld, A. (1985) *Mathematical problem solving*. New York: Academic.

Schoenfeld, A. (2002). Research methods in (mathematics) education. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 435–487). New Jersey: Lawrence Erlbaum.

Schubauer-Leoni, M. L., & Grossen, M. (1993). Negotiating the meaning of questions in didactic and experimental contracts. *European Journal of Psychology of Education, 8*(4), 451–471.

Van Someren, M. Y., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modeling cognitive processes*. London: Academic.