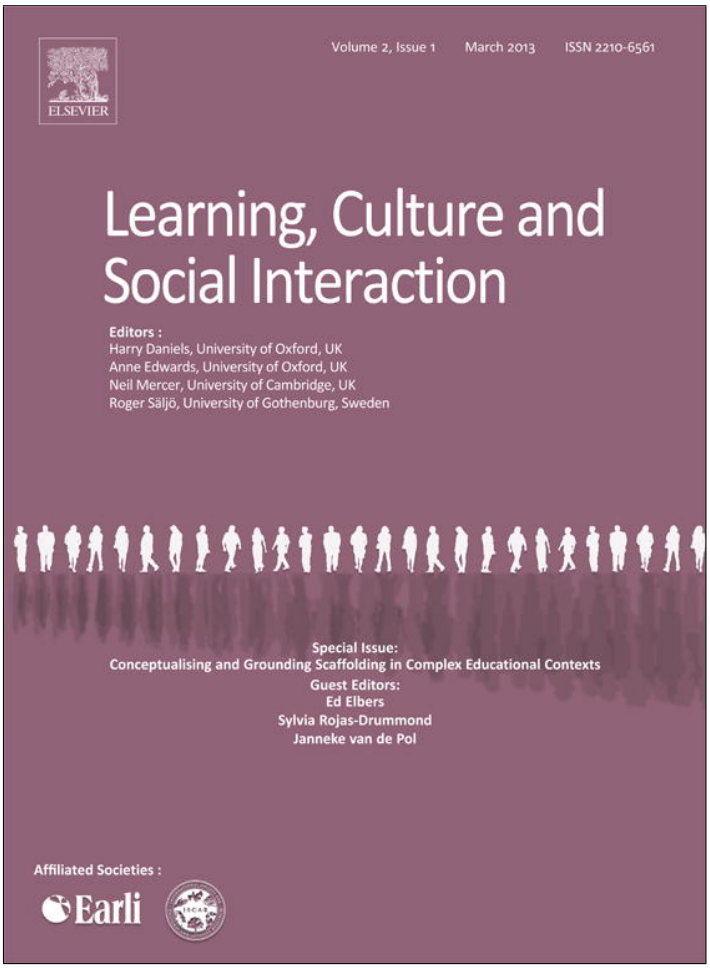

Learning, Culture and Social Interaction 2 (2013) 22–31

Contents lists available at SciVerse ScienceDirect

Learning, Culture and Social Interaction

journal homepage: www.elsevier.com/locate/lcsi

What counts as evidence for the long-term realisation of whole-class scaffolding?

Jantien Smit *, Dolly van Eerde

Freudenthal Institute, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 4 October 2012
Accepted 21 December 2012
Available online 8 February 2013

Keywords:
Handover to independence
Responsiveness
Scaffolding
Whole-class interaction

ABSTRACT

To promote whole-class scaffolding of mathematical language, a teacher was encouraged to employ a repertoire of seven strategies (e.g., reformulating) in a multilingual primary classroom (22 pupils; aged 10–12). This paper investigates whether the enactment of these strategies has led to long-term whole-class scaffolding as identifiable by its key characteristics: diagnosis, responsiveness and handover. Comparison of pupils' pre- and post-test scores on three linguistic key elements all yielded statistically significant differences with large effect sizes, thus confirmed handover. A statistically significant shift from high-support to low-support strategies revealed responsiveness to pupils' levels over nine lessons. A qualitative analysis showed interrelatedness of performed strategies and scaffolding characteristics (e.g., diagnosis). The results provide empirical evidence of the long-term realisation of whole-class scaffolding.

1. Introduction

The metaphor of scaffolding originates from the context of problem solving to conceptualise adults' supportive role in dyadic adult–child interaction (Wood, Bruner, & Ross, 1976). Scaffolding nowadays refers to the temporary, responsive support that helps a pupil perform a task he cannot yet complete by himself and that is intended to bring the pupil gradually to a state of competence in which he can complete a similar task independently (Jadallah et al., 2011). Ever since its introduction, there has been a growing body of research investigating how to describe, enact and analyse scaffolding (Lin et al., 2012; Van de Pol, Volman, & Beishuizen, 2010). Along with its increasing popularity, the concept has been extended to a variety of settings other than one-to-one teacher–student interaction. One complex educational context for which it is advocated is the whole-class setting (e.g., Hogan & Pressley, 1997). Research on whole-class scaffolding, however, faces conceptual and methodological challenges.

1.1. Conceptual challenges

Concept extension can lead to overgeneralisation and loose use of the concept if not thought through carefully. Overgeneralisation is involved when the concept is stretched so far that almost any support in classroom interaction (e.g., Meyer & Turner, 2002), or even classroom organisation, artefacts or sequencing (e.g., Anghileri, 2006) is referred to as scaffolding. This trend has been widely criticised over the last decade (e.g., Puntambekar & Hübscher, 2005). Loose use is the case when the concept of scaffolding is used in classroom studies without explicit reference to its key characteristics (as is the case in Jadallah et al., 2011).

In response to these conceptual challenges faced in research on whole-class scaffolding, we have proposed a conceptualisation of whole-class scaffolding that is both theoretically informed and empirically grounded (Smit, Van Eerde, & Bakker, in press). Based on the scaffolding literature, three key characteristics of whole-class scaffolding are distinguished: diagnosis, responsiveness (the adaptive core of the support provided) and handover to independence. These characteristics serve to identify teaching in

* Corresponding author. Tel.: +31 30 253 5662.
E-mail addresses: J.Smit1@uu.nl (J. Smit), H.A.A.vanEerde@uu.nl (D. van Eerde).

2210-6561/$ – see front matter
http://dx.doi.org/10.1016/j.lcsi.2012.12.006

whole-class settings as whole-class scaffolding. Based on a teaching experiment in which whole-class scaffolding was enacted, we formulated three features of whole-class scaffolding that capture its nature, and that distinguish whole-class scaffolding from small-scale scaffolding as conceptualised in the literature up to now. (1) In the first place, we metaphorically distinguished between *online* and *offline* enactment of key characteristics, during respectively outside whole-class interaction (layered nature). Analysing pupils' work between lessons is an example of offline diagnosis. An example of offline responsiveness is adapting instructional activities between lessons. Offline handover to independence is fostered in the design of lessons. (2) In the second place, key characteristics appeared to occur not in one single episode of classroom interaction, but distributed over several episodes (distributed nature). For example, diagnosing during a lesson can lead to responsiveness in a subsequent lesson. (3) Lastly, pupils' increase in independence has been conceptualised as the cumulative effect of several instances of diagnoses and responsiveness over time (cumulative nature). Our conceptualisation aimed to address the aforementioned conceptual challenges.

1.2. Methodological challenges

Several studies have yielded frameworks or categorisations of scaffolding focusing on teacher actions, but without studying effects on pupils (e.g., Nathan & Knuth, 2003). Scaffolding, however, is relational by nature. If studies do include pupils' progress, it is often unclear how teachers' strategies are related to pupils' development. In other words, teachers' adaptation of support to pupils' level of knowledge or ability needed for pupils' learning is rarely investigated in such studies. Yet responsiveness (also referred to as contingency) has been argued to be the most distinctive characteristic of scaffolding (e.g., Van de Pol et al., in press). Moreover, explicit attention to teachers' diagnosing pupils' levels is also rare (an exception is Van de Pol, Volman, & Beishuizen, 2011). If we intend to do justice to the relational nature (between teacher and pupils) of whole-class scaffolding (e.g., Rasku-Puttonen, Eteläpelto, Arvaja, & Häkkinen, 2003) we face the methodological challenge of analysing diagnosis, responsiveness and handover also in relation to each other (cf. Mercer & Fisher, 1992). We only assert whole-class scaffolding if a teacher diagnoses pupils' linguistic levels and responsively performs strategies, with indications of handover as a result.

An implication of this choice is that units of analysis larger than brief fragments of classroom interaction are typically needed, as "the process of teaching and learning in school has a natural long-term trajectory" (Mercer, 2008, p. 33). As a consequence we need to analyse classroom interaction as "linked chains of interaction" (Scott, Mortimer, & Aguiar, 2006, p. 610).

1.3. The current study

In a design-based research project we have developed a set of strategies for scaffolding the development of subject-specific language needed for mathematical learning, in an attempt to enact whole-class scaffolding (Smit & Van Eerde, 2011). In the current paper we deal with methodological challenges in the context of scaffolding in whole-class settings. Our aim is to investigate whether the enactment of strategies intended to scaffold the development of subject-specific language in an upper primary mathematics classroom has led to whole-class scaffolding as identifiable by its key characteristics, taking into account the aforementioned long-term nature of whole-class scaffolding. To fulfil this aim we address three research questions in response to the aforementioned methodological challenges.

1. To what extent did handover to independence take place?
   To determine whether whole-class scaffolding has taken place, we need to find evidence of a process of handover to independence. Unlike most scaffolding studies that focus on the development of reasoning skills or content knowledge (e.g., Nathan & Kim, 2009; Speer & Wagner, 2009), our focus of independence is on subject-specific language development. Scaffolding language has been advocated particularly in multilingual classrooms, and has been argued to be a prerequisite for second language learners' successful participation at school. A reason put forward is that the development of subject-specific language is intertwined with conceptual development in a particular school subject (e.g., Gibbons, 2002). In our case, a process of handover concerned pupils' development of subject-specific language needed for reasoning about line graphs.
   Finding evidence of handover to independence only, however, does not sufficiently underpin the assertion of scaffolding. We therefore also investigate the characteristic of responsiveness in relation to performed teacher strategies:
2. What evidence of responsiveness can be identified in the teacher's enactment of strategies for promoting language development?
   In the literature on scaffolding few studies explicitly examine responsiveness as part of classroom scaffolding. Those that do tend to have a local character in the sense that they concern teacher–student or teacher–small group interaction rather than whole-class teaching. Moreover, small units of analysis are used (varying from a single utterance to three-turn sequences or brief interaction fragments), implying that only brief fragments of interaction are investigated to identify responsiveness or contingency (e.g., Van de Pol et al., in press). Given the interactive nature of scaffolding it is defendable to study a teacher's adaptive behaviour at such micro-levels. As argued before, however, many topics of learning, such as (second) language development, take longer periods of time to pursue (Dixon et al., 2012; Mercer, 2008). In line with our conceptualisation of whole-class scaffolding we think that analysis of both pupils' language development and responsiveness of performed strategies thus also needs to be framed within a larger time scale.
   Once we have gained evidence of handover and responsiveness, the question arises how diagnosis occurred, and we analyse this in relation to the other scaffolding characteristics to show how the strategies and characteristics are related. To this end we need fine-grained analysis. Hence our third research question:
3. How are the performed strategies and characteristics (diagnosis, responsiveness and handover) of whole-class scaffolding related over time?

As Mercer (2008) noted, few researchers have taken up the challenge to analyse episodes of classroom talk as the aforementioned linked chains of interaction. This lack of research can be explained by the theoretical and methodological challenges posed by the examination of the relationship between time, talk and learning in classrooms (as argued by Littleton, 1999). One challenge is how a class of pupils can be diagnosed: Should all individual pupils be assessed or does the online diagnosis of a sample suffice to get a sense of the group zone of proximal development (Smit et al., in press)? In fact, offline diagnosis (e.g., diagnosing student work between lessons) allows a teacher to be responsive to a whole class. Therefore such offline enactment of key characteristics is also included in our analysis.

To understand how pupils' language development unfolds over time in the context of whole-class scaffolding, we need 'story threads' to uncover crucial moments in pupils' performance (during lessons) and in the teacher's performance (during *and* between lessons). A story thread can be seen as a "chain of events" (Nathan & Kim, 2009, p. 113) that illustrates a causal field — a multitude of factors in relation to each other. As an inevitable consequence of the complexity such qualitative analysis brings along, we present only one story thread in this paper, focusing on one aspect of the language to be developed. We could have elaborated multiple story threads, but this one is representative in the sense that it concisely illustrates the layered, distributed and cumulative nature of whole-class scaffolding (another example is presented in Smit et al., in press).

2. Methods

2.1. Research setting and participants

This study is part of a design-based research project in which three teaching experiments were carried out. The overall aim of the project is to investigate how a teacher in multilingual classrooms can scaffold the development of language needed for mathematical reasoning about line graphs. Design research as envisioned in educational research (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003) was a suitable methodological approach to accomplish this, as it aims at developing innovative instructional activities *and* theories. Its highly interventionist nature and cyclic character allowed us to theoretically develop the concept of whole-class scaffolding while at the same time promoting and developing its enactment in classroom situations. In designing lessons, we focused on both mathematical knowledge development and (second) language development. We capitalised on the latter in the following two ways.

In the first place, we used the teaching and learning cycle (e.g., Gibbons, 2002) as a design heuristic. This cycle consists of a series of four stages in which a particular academic text genre needed for a particular school subject is introduced, modelled, jointly practised, and eventually individually performed by the pupils. Support is thus withdrawn once pupils gain in proficiency (offline handover to independence). The underlying idea is that pupils need gradual induction into the subject-specific academic language involved in a genre. In our case, the genre to be developed in pupils was the *interpretative description of a line graph* (Smit, Van Eerde, Kuijpers, & Bakker, submitted for publication). A text in this genre both describes the course of the line graph (e.g., "the graph descends gradually") and interprets its meaning (e.g., "he slowly loses weight"). The genre includes both linguistic features (e.g., the use of subject-specific words) and structure features (e.g., describing each segment of the graph in terms of reality and in terms of the graph's course).

In the second place, we developed a conceptual framework of strategies for scaffolding language (see Table 1). This framework, constructed to both promote and evaluate the teacher's scaffolding performance, was refined in three cycles of design research. Its construction was firstly informed by scaffolding literature, which is for instance reflected in the potentially responsive nature of all strategies. Secondly, its construction was informed by the empirical data, as these informed us on the extent to which strategies seemed effective for pupils' language development.

For this paper, we only use data from the third teaching experiment, consisting of nine lessons. Because whole-class scaffolding was promoted, classroom interaction predominantly took place in whole-class settings, implying that one-to-one and small group discussions were much less frequent. The experiment was carried out in a combined grade 5–6 class of a suburban primary school (age 10 – 12). Lessons were given weekly and lasted 60 to 70 min each. The number of pupils speaking Dutch as a second language was 19 out of 22, the majority being second or third generation Moroccan and Turkish pupils, who performed rather weakly on a standardised test for language (CITO assessment). The teacher, who had also participated in the two previous teaching experiments, had seventeen years of experience in primary education, partly in multilingual classrooms.

Table 1
Strategies for scaffolding language and examples for each strategy.

1	Reformulate pupils' utterances (spoken or written) into more academic wording	[In response to *the graph goes higher and higher up*:] *Yes, the graph does rise steeply.*
2	Ask pupils to be more precise in spoken language or to improve their spoken language	*What do you mean by 'it'?*
3	Repeat correct pupil utterances	*Yes, the graph does descend slowly.*
4	Refer to features of the text type (interpretative description of a line graph)	*Into how many segments can we split the graph?*
5	Use gestures or drawings to support verbal reasoning	E.g., gesturing a horizontal axis when discussing this concept
6	Remind pupils (by gesturing or verbally) to use a designed scaffold (i.e. word list or writing plan) as a supporting material	*Look, the word you are looking for is written down here.*
7	Ask pupils how written text can be produced or improved	*How can we rewrite this in more mathematical language?*

2.2. Instruments

We used a pre- and a post-test to determine handover to independence. These tests consisted of respectively nine and eight items and lasted up to 50 min (pre-test) and 45 min (post-test). There were eight days between the last lesson and the post-test. The selection of three test items used for answering the first research question was informed by our choice of the three linguistic key elements needed for reasoning about line graphs, involving various estimated learning burdens (cf. Nation, 2001):

1. A group of graph-related words: horizontal axis, vertical axis and title.
2. A conceptually difficult key word in the domain of line graphs, used to reason about issues of steepness: 'gradually'.
3. Temporal prepositions: words referring to a particular moment in time, represented by a point in the line graph, for instance *at* 6 o'clock; word combinations referring to a particular period in time, represented by a segment of the line graph, for instance *between* 6 *and* 8 (particularly difficult for second language learners and key to reasoning adequately in the domain of graphs; Chodorow, Gamon, & Tetreault, 2010).

The test items were constructed so as to gain insight into both receptive and productive word knowledge, based on the underlying assumption that both kinds of vocabulary development should be included in tests as they indicate different degrees of proficiency in a second language (Nation, 2001). Receptive knowledge of the graph-related words was measured in a test item in which pupils were asked to mark the horizontal axis, the vertical axis and the title of a line graph. In a subsequent test item pupils were asked to write down the meaning of a number of subject-specific words. This item was used to investigate pupils' knowledge of 'gradually'. In the third test item pupils were asked to describe a line graph (without support) for a dual purpose: analysing pupils' productive knowledge of both subject-specific language (including 'gradually') and analysing pupils' productive knowledge of temporal prepositions (see Fig. 1). These types of productive knowledge are analysed separately (Smit et al., submitted for publication).

2.3. Data collection

Apart from the pre- and post-test scores of all 22 pupils, data collection included video and audio recordings of all lessons; all recordings were transcribed verbatim. Furthermore, data collection consisted of pupils' written work, the teacher's weekly written reflections using a reflection format we gave her, audio recordings of video-stimulated recall interviews (SRIs) one of the researchers held with the teacher between lessons (Smit & Van Eerde, 2011), as well as reports on the SRIs written by the researcher. The underlying idea of SRIs is to encourage teachers to think aloud while observing videos of their own lessons, thus stimulating them to relive the lesson (Meijer, Zanting, & Verloop, 2002). At the end of each SRI we discussed the teacher's written reflection, which included questions on her diagnoses, the responsiveness of her strategy performance and her observations of pupils' progress. The characteristics of whole-class scaffolding were also central to these weekly discussions, on the basis of which the teacher formulated scaffolding intentions for the next lesson.

2.4. Data analysis

2.4.1. To what extent did handover to independence take place?

To compare pre- and post-test scores on the aforementioned items, we first investigated whether the scores were distributed normally, using the Kolmogorov–Smirnov normality test (in SPSS), to see if we could use the paired-samples t-test. For the first three comparisons, this was not the case ($p = <.001$), so we used the non-parametric Wilcoxon's signed rank test. In the last comparison, on temporal prepositions, we could assume normality ($p = .11$) but for better comparison we used Wilcoxon in all cases along with the associated effect sizes (r).

Fig. 1. Test item for investigating both pupils' productive knowledge of subject-specific language and their productive knowledge of temporal prepositions. The accompanying text was: Describe how Auntie Mary's weight changes and how you can tell from the graph.

Table 2
Coding scheme for identifying high-support (H) and low-support (L) strategies for scaffolding language.

Code	Strategy for scaffolding language
RefExt (H)	Reformulating or extending pupils' spoken or written utterances
LingFea (H)	Explicitly referring to or reminding of linguistic features (e.g., subject-specific words or temporal prepositions), or doing so implicitly by referring to or pointing at the word list, or by referring explicitly to supportive gestures
StrucFea (H)	Explicitly referring to or reminding of structure features (e.g., the use of a specific type of language such as mathematical language)
AskImp (L)	Asking pupils to improve language (e.g., asking for more precise language) or to elaborate their utterance
RepCor (L)	Literally repeating correct pupils' utterances or making explicit the good quality of pupils' utterances or gestures
IndPro (L)	Asking for or explicitly encouraging pupils' to independently produce spoken or written language

Table 3
Example of instructions for coding.

Code	Coding instruction for each strategy
RefExt	1) Attribute code for reformulations: Pupil: The graph goes descending. Teacher: Indeed, the graph descends. [correcting grammatical error]
	2) Attribute code for extensions: Teacher: So what do I need to do now? Pupil: Coordinate system. Teacher: Now I am going to draw the coordinate system.

2.4.2. What evidence of responsiveness can be identified?

Our hypothesis was that responsiveness over time should be visible in a shift from high-support to low-support strategies. High-support strategies involve a high level of teacher support and low-support strategies involve a low level of teacher support, thus posing bigger demands on pupils' independence. To examine such a shift a coding scheme with instructions was developed in several rounds (see Tables 2 and 3). The scheme was based on the aforementioned conceptual framework (Table 1), although we slightly regrouped the categories. As a unit of analysis we used the smallest meaningful transcript segment in which the teacher's endeavour to scaffold language could be identified (cf. Rasku-Puttonen et al., 2003). This could be a word, e.g., "horizontal" (when repeating a correct pupil utterance), a sentence, e.g., "What did we call that again?" (when asking for more precise language), or a few sentences which taken together form a strategy for scaffolding language.

To identify a possible shift, we decided to contrast the earliest and the latest segments in which each of the three aforementioned linguistic key elements was addressed. To this end we needed to balance between two requirements. First, for a contrast we needed a sufficiently large gap between segments to be coded. Second, for measuring the interrater reliability of the coding scheme we needed at least $2n^2$ segments (in our case 72), where n is the number of codes (Cicchetti, 1976). The balance was found in analysing a maximum of two early and two late lessons in which strategies were performed. This resulted in a total number of 80 segments to be coded (see Fig. 2 for the distribution of segments over lessons). A second researcher independently coded the 80 segments, resulting in 74 agreements in coding (92.5%) and Cohen's kappa of .90, which indicates a very high interrater reliability. The resulting data were analysed with Pearson's chi-squared test.

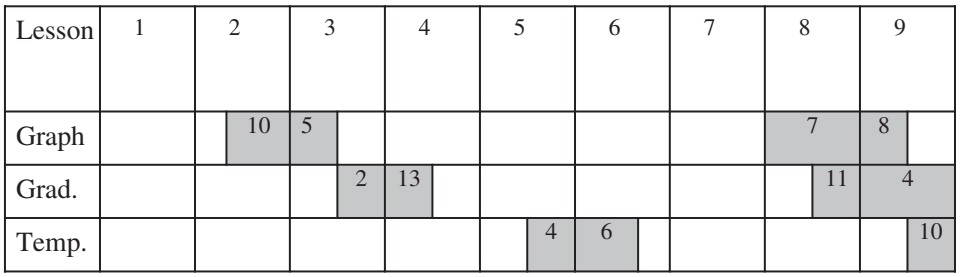

Fig. 2. Distribution of segments to be coded: 30 concerning graph-related words (Graph), 30 concerning the use of 'gradually' (Grad.) and 20 concerning the use of temporal prepositions (Temp.).

Table 4
Pre- and post-test scores and results of Wilcoxon signed rank test and effect sizes.

	Topic	Pre-test mean (SD)	Post-test mean (SD)	Max score	Statistic *z*	Effect sizes (r)	Ranks (positive, tie, negative)	*p*
1	Graph-related words (total)	0.59 (.73)	2.36 (1.09)	3	3.70	.79	17 +; 4 ties; 1 −	<.001
	– Title	.50	.91	1				
	– Horizontal axis	.045	.73	1				
	– Vertical axis	.045	.73	1				
2	Gradually	0.0	.73	1	4.00	.85	16 +; 6 ties; 0 −	<.001
3a	Subject-specific	.16 (.47)	2.68 (1.39)	4	3.87	.83	19 +; 3 ties; 0 −	<.001
3b	Temporal prepositions	2.20 (.55)	3.84 (1.51)	6	3.53	.53	16+; 6 ties; 0 −	<.001

2.4.3. How are the performed strategies and characteristics of whole-class scaffolding related over time?

To answer the third research question we focus on a group of graph-related words that was introduced early in the lesson series (cf. Fig. 2): horizontal axis, vertical axis, coordinate system, graph and title. We illustrate the journey of these words towards independence throughout the lesson series. In chronological order, we identified and coded all segments in the lesson transcripts in which the teacher performed strategies concerning these graph-related words. We read the transcripts around these segments in search of indicators of pupils' progress (or decline) and the teacher's responsiveness (or lack of responsiveness) to pupils. Subsequently, we identified all teacher comments in relation to these graph-related words during the SRIs that could be interpreted as diagnosis, or an intention to responsiveness or handover to independence. From the chronological list of both transcript fragments and teacher comments we selected those elements that signified change or improvement (there were very few examples of decline). These elements form the basis for the story thread.

3. Results

3.1. Research question 1: handover to independence

The comparisons of pre- and post-test scores all yielded statistically significant differences (see Table 4). The effect sizes are large ($r = z/\sqrt{N} = .5$ is considered large, N is the number of measurements, $2^*22 = 44$). For readers who are used to effect sizes measured with Cohen's *d*, we also report $d = 1.13$ on temporal prepositions (0.8 is considered large). This suggests that pupils have become much more independent on these linguistic aspects of the genre of reasoning about line graphs. Thus handover to independence occurred to a large extent.

3.2. Research question 2: responsiveness

Table 5 shows the frequencies of strategies performed per group. The frequencies suggest a shift from high-support to low-support strategies. This shift is indeed statistically significant, χ^2 (1, $N = 80$) = 22.75, $p < .001$. This shows evidence of the teacher's responsiveness to pupils' levels of language proficiency in her enactment of language promoting strategies.

3.3. Research question 3: story thread about graph-related words

Having given evidence of handover to independence and responsiveness by means of quantitative data from the beginning and end of the lesson series, we now use qualitative data as the basis for a connecting story thread that shows how the teacher strategies are linked to the scaffolding characteristics of diagnosis, responsiveness and handover. We use abbreviations to refer to lessons and strategies; for example 'L2 LingFea' means lesson 2, and performance of the LingFea strategy (see Table 2) concerning linguistic features.

In lesson 2, the teacher introduces the term 'horizontal axis':

> What is 'horizontal axis'? Is that this one (moves her hand vertically) or is this one (moves her hand horizontally). What is the horizon? I always remind myself: what is the horizon. Who knows? The horizon. Think of the sea. [L2 LingFea].

Table 5
Frequencies of high- and low-support strategies.

	First period	Last period	Total
High-support strategies	34	13	47
Low-support strategies	6	27	33
Total	40	40	80

The teacher performs the high-support LingFea strategy a few more times, i.e., she centralizes the newly introduced concept by referring to phenomena pupils know (the horizon at sunset). She also introduces the related words 'vertical axis' and 'coordinate system'. In line with theories on second language development (Dixon et al., 2012) the teacher returns to the introduced words several times. We consider these instances online responsive to pupils' linguistic levels: only one pupil (out of 22) demonstrated receptive word knowledge of horizontal axis and vertical axis in the pre-test (offline diagnosis).

In the second half of lesson 2 the teacher starts extending pupils' utterances by using the graph-related words in full sentences or by adding the definite article (both RefExt strategy). The definite article is particularly difficult for second language learners of Dutch because this language uses two different articles (one for masculine and feminine words, and one for neutral words). Both the extension and reformulation of pupils' utterances can be considered as online responsively performed strategies in the case of emerging word knowledge.

The teacher also encourages pupils to use the word list on the wall as supportive material (LingFea strategy). This list is used both to write down newly introduced subject-specific words, and to return to these words (verbally or by gesturing) when required in whole-class mathematics discussions, as the following fragment illustrates.

Teacher	Step one. What am I going to do?
Bela	At the coordinate system…
Teacher	Very good, isn't it? You may not have seen it, but I saw her eyes. Rrrts went her eyes (teacher points from Bela to the word list on the wall) [L2 LingFea]. Coordinate system. That is very smart, because it takes some time for that word to sink in. (…) Does it matter which one [axis] I draw first? No, but we want to practise using those words.

The fragment also exemplifies how the teacher establishes a norm concerning the need to employ subject-specific words for discussing mathematical content. Establishing such norms contributes to pupils' awareness of the need to learn subject-specific language for participating in mathematical discussions. In subsequent interaction fragments in lesson 2 several pupils do start using the graph-related words (supported by the word list). One of the few pupils whose native language is Dutch, is the first one to include horizontal axis in a full sentence: "Then you will, you write centimetres along the horizontal axis."

In the SRI after lesson 2 the teacher states from memory: "I presume five pupils to have added the concept 'horizontal axis' and 'vertical axis' to their vocabulary." We infer online diagnosis from this statement as it indicates that the teacher has tacitly diagnosed pupils' word knowledge. During this SRI the teacher expresses the presumption that pupils will need less extensive linguistic support in the next lesson, which indicates the start of a process of handover: "Next time only gesturing will be sufficient support, or at least I will spend fewer words explaining the meaning [of horizontal axis]."

What struck us from the analysis was the importance of intentions as connecting diagnosis to future responsiveness or handover. The first crucial intention, expressed in the same SRI, was to support pupils in actively using the targeted language, based on the online diagnosis of pupils' language use:

> And for Abdul…he really speaks what I call 'street language'. I just need to encourage him to use this [subject-specific] language correctly a few times. He needs to experience what it is to talk like this; he needs to experience that he can actually produce this kind of language.

The second intention, also based on a tacit online diagnosis, concerns the following general aspects: "I want to focus more on pupils completing their sentences, as well as on those pupils who haven't yet verbally contributed."

During lesson 3 the teacher returns to the previously introduced words and notices many pupils are now able to gesture in the right directions of both axes. She makes this progression explicit to pupils (online handover; RepCor strategy):

Teacher	What was it again, "horizontal"?
Abdul	(Abdul moves his hands to and fro in horizontal direction)
Teacher	You go like this already! (Teacher makes the same movement as Abdul.) Yes. Good, isn't it? Indeed. That was horizontal. [L3 RepCor].

During the SRI after lesson 3 the teacher articulates this process as follows: "It goes exactly as I predicted earlier: pupils are taking over the gesture for horizontal axis and less words are needed to explain the concept." The teacher seems to have internalised the promoted new way of teaching, as she perceives such statements as on-the-spot decisions and thoughts of her own. The following statement out of the same SRI also indicates a process of handing over, as the teacher infers from her tacit online diagnoses: "Horizontal axis and vertical axis are starting to sink in, but these concepts are not yet known to everybody, so more practice is needed here."

During lesson four the teacher again pays explicit attention to the concept of horizontal axis (LingFea strategy) due to one pupil mixing up the axes. Soon after that a Dutch language learning pupil, Abdul, productively uses the concept by himself in an adequate way. His development from gesturing (lesson 3) to producing language (lesson 4) indicates increased linguistic independence:

Teacher	What did we do next, Abdul?
Abdul	Weeks at, the weeks along, writing the weeks along the horizontal axis.
Teacher	Indeed. [L4 RepCor]
Abdul	And the time.
Teacher	We have written time in weeks along the horizontal axis. [L4 RefExt] And?
Abdul	And along the vertical axis we had to write centimetres.

Later in lesson 4 the teacher asks another pupil to independently produce a sentence (IndPro strategy). This is the first time she uses this low-support strategy (see following fragment). We consider this interactively responsive to the spontaneous word use by Abdul earlier in the same lesson.

Yassin	Horizontal axis (points to the horizontal axis).
Teacher	Can you put that in a sentence? [L4 IndPro]

During the SRI after lesson 4 the teacher does seem to have diagnosed pupils' progressing word knowledge: "Pupils' knowledge of words has increased, I concluded from the round I conducted." This online diagnosis is confirmed by the teacher's offline diagnosis based on pupils' linguistic levels in written work: for example, 16 out of 22 pupils have acquired receptive word knowledge of horizontal axis. As a consequence, we decide to focus our attention on other words and linguistic aspects still in need of scaffolding.

Despite our shift of attention, pupils' language development concerning the graph-related words continues to evolve: spontaneous use of the concepts happens more frequently and pupils more often use these concepts in longer phrases. In line with this development, Table 6 shows a relative increase of low-support strategies (RepCor, AskImp, IndPro) and a decrease of high-support strategies (LingFea, RefExt). Towards the end of the lesson series, the teacher also refers to pupils' gained independency several times, as for instance in: "For that we know proper mathematical language" (L7 AskImp). The following fragment from lesson 7 exemplifies this stage in the process of handover, concerning both pupils' and the teacher's development.

Semi	He weighs 85 kg.
Teacher	He weighs 85 kg [L7 RepCor]. OK. Hamid, how can he tell?
Hamid	From the vertical axis.
Teacher	From the vertical axis, yes [L7 RepCor]. Along the vertical axis we see the data on uncle Kees' weight [L7 RefExt]. And what informs him about his age? Youness.
Youness	Down below. (…)
Teacher	And down below, for that we know proper mathematical language [L7 AskImp]. Oussana.
Oussana	Horizontal axis.
Teacher	Now put that in a sentence [L7 IndPro]. That's just a word, 'horizontal axis' is just a word on its own like that.
Oussana	Along the horizontal axis it says age in years.
Teacher	That's a beautiful sentence, isn't it [L7 RepCor]?

The following remark during the SRI after lesson 7 points to online diagnosis, and to adjusting an earlier made diagnosis: "I hadn't expected they would speak in full sentences so well."

With this story thread we hope to have given an impression of the qualitative relationships between teacher strategies and scaffolding characteristics over time, hence the texture of whole-class scaffolding. First, the story thread points to how diagnosis is related to responsively using strategies: they are connected by intentions. Secondly, the story thread exemplifies possible relationships between online and offline diagnosis: confirmation or adjustment. Confirmation was at stake when offline diagnosis endorsed what the teacher had diagnosed online on the basis of a small sample of students. Adjustment was necessary when a new diagnosis refuted one made earlier. A third point we want to make is that this story thread, in particular Table 6, confirms a relationship that we already found in the answer to research question 2: the shift from high- to low-support strategies indicates responsiveness over time. In sum, the cumulative online enactment (e.g., responsive performance of strategies) and offline

Table 6
Frequencies of strategies for the scaffolding of graph-related words performed in each lesson.

Lesson	2	3	4	5	6	7	8	9
Strategy								
RefExt	3	1	2	2	2		2	
LingFea	6	7	7	2		1	3	1
StrucFea								
AskImp			1	8	2	2	1	2
RepCor	1	2	4	4		2	5	5
IndPro				1			2	
Total	10	11	14	17	4	5	13	8

Note: The StrucFea strategy was not enacted for this particular linguistic key element.

enactment (e.g., diagnosis) of key characteristics over time eventually led to handover to independence. This story thread thus illustrates the layered, distributed and cumulative nature of whole-class scaffolding.

4. Discussion

The present study aimed to investigate whether the enactment of strategies intended to scaffold the development of a subject-specific genre did lead to whole-class scaffolding as identifiable by its key characteristics (handover, responsiveness and diagnosis). We employed a long-term conceptualisation and analysis of whole-class scaffolding. In answer to the first research question, we can conclude that pupils have become significantly more independent on the measured linguistic key elements of the genre needed for reasoning about line graphs. This points to a long-term process of handover to independence. The second research question's answer can be found in a statistically significant shift from high-support to low-support strategies. This shift indicates that the teacher adjusted her performance of strategies to pupils' levels of language proficiency. Her strategy use was therefore responsive over the long term. In answer to the third research question, we can conclude that performed strategies, the three key characteristics of whole-class scaffolding (diagnosis, responsiveness, handover) and pupils' progress appeared to be closely linked over time. These answers provide empirical evidence of the realisation of whole-class scaffolding by means of strategies for promoting pupils' language development.

Although several researchers have indicated the importance of linking talk and learning over time (e.g., Mercer, 2008), such long-term analyses have rarely been conducted in scaffolding research (an exception is Rasku-Puttonen et al., 2003). We employed mixed methods to benefit from the strengths of both approaches and to generate evidence of different kinds (cf. Mercer, 2010). The quantitative methods used allowed for contrasting numerical comparisons between pupils' proficiency levels as well as between teacher strategies performed in early and late lessons. To give meaning to these quantitative results, a qualitative analysis provided a detailed examination of what the process of whole-class scaffolding actually looked like, linking strategies, key characteristics and pupils' progress over time. Hence the methodological significance of this study lies in yielding methods for analysing whole-class scaffolding that is intended to foster long-term learning processes. We presume that such methods can be used for analysing whole-class scaffolding in other domains as well because they do not seem to depend on domain-specific knowledge.

A first limitation concerns the rather confined measurement of handover to independence. Second language theorists, for instance, may suggest to include pupils' ability to connect newly acquired words (e.g., gradually) to words already present in the learner's lexicon in the analysis (cf. Ellis, 1997). From a sociocultural perspective one may argue for measuring independence in a broader sense (cf. Lockhorst, Wubbels, & Van Oers, 2010), for instance as a manifestation of an emerging language community. Research on such a language community could include analysis of classroom discourse: how and to what extent does pupils' use of subject-specific language evolve over time? Furthermore, it could include measuring pupils' own initiative in whole-class settings, or pupils' scaffolding of each other (cf., Fernandez, Wegerif, Mercer, & Rojas-Drummond, 2001).

A second limitation concerns the representativeness of our story thread. The story thread provided an answer to the third research question in that it illustrated how performed strategies were related to online and offline enactment of key characteristics over time, in this particular study. However, further research is needed to verify whether the relationship patterns found here (e.g., between online and offline diagnosis) also occur in other whole-class scaffolding situations. Such research can also deepen our understanding of the layered, distributed and cumulative nature of whole-class scaffolding.

A question that may rise is whether our study provides a warrant for suggesting that teacher strategies have indeed been the main cause of pupils' increased independence. Inevitably, there have been more factors at stake that may have contributed to pupils' progress (e.g., the teacher providing occasional individual support). Instead of asserting specific causal relations between the enactment of individual strategies and effects on pupils' development, we prefer to view the evidence of whole-class scaffolding as pointing to a causal field. We think it is highly plausible that strategies performed in our teaching experiment have been vital in this causal field. A strong indication is pupils' better appropriation of genre elements after the teacher's enactment of strategies, as visible in the story thread. As Jadallah et al. (2011, p. 226) argue, such evidence is not decisive but the likelihood that pupils would say what they say is “vanishingly small” without the teacher's influence.

In sum, we think we have provided empirical evidence for the long-term realisation of whole-class scaffolding as a result of performing particular strategies. Because these are not domain-specific, we think that they can also be applied for the whole-class scaffolding of other subject-specific genres required throughout the curriculum.

Acknowledgements

We thank Maggie Kuijpers for conducting one part of the analysis and the teacher for her committed participation in the teaching experiment. Jan van Maanen, Koeno Gravemeijer, Arthur Bakker and the guest editors provided us with insightful comments. We thank Nathalie Kuijpers for editing the English text.

References

Anghileri, J. (2006). Scaffolding practices that enhance mathematics learning. *Journal of Mathematics Teacher Education*, *9*, 33–52.

Chodorow, M., Gamon, M., & Tetreault, T. (2010). The utility of article and preposition error correction systems for English language learners: Feedback and assessment. *Language Testing*, *27*(3), 419–436.

Cicchetti, D. V. (1976). Assessing inter-rater reliability for rating scales: Resolving some basic issues. *The British Journal of Psychiatry*, *129*, 452–456.
Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*, 9–13.
Dixon, L. Q., Zhao, J., Shin, J. -Y., Wu, S., Su, J. -H., Burgess-Brigham, R., et al. (2012). What we know about second language acquisition: A synthesis from four perspectives. *Review of Educational Research*, *82*(1), 5–60.
Ellis, R. (1997). *Second language acquisition.* New York: Oxford University Press.
Fernandez, M., Wegerif, R., Mercer, N., & Rojas-Drummond, S. (2001). Re-conceptualizing "scaffolding" and the zone of proximal development in the context of symmetrical collaborative learning. *Journal of Classroom Interaction*, *36*(2), 40–54.
Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom.* Portsmouth, NH: Heinemann.
Hogan, K., & Pressley, M. (1997). Scaffolding scientific competencies within classroom communities of inquiry. In K. Hogan, & M. Pressley (Eds.), *Scaffolding student learning* (pp. 74–107). Cambridge, MA: Brookline Books.
Jadallah, M., Anderson, R. C., Nguyen-Janiel, K., Miller, B. W., Kim, I. -H., Kuo, L. -J., et al. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussion. *American Educational Research Journal*, *48*(1), 194–230.
Lin, T. -C., Hsu, Y. -S., Lin, S. -S., Changlai, M. -L., Yang, K. -Y., & Lai, T. -L. (2012). A review of empirical evidence on scaffolding for science education. *International Journal of Science and Mathematics Education*, *10*, 437–455.
Littleton, K. (1999). Productivity through interaction: An overview. In K. Littleton, & P. Light (Eds.), *Learning with computers: Analyzing productive interaction* (pp. 179–194). London: Routledge.
Lockhorst, D., Wubbels, T., & Van Oers, B. (2010). Educational dialogues and the fostering of pupils' independence: The practices of two teachers. *Journal of Curriculum Studies*, *42*(1), 99–121.
Meijer, P. C., Zanting, A., & Verloop, N. (2002). How can student teachers elicit experienced teachers' practical knowledge? Tools, suggestions and significance. *Journal of Teacher Education*, *53*, 406–419.
Mercer, N. (2008). The seeds of time: Why classroom dialogue needs a temporal analysis. *The Journal of the Learning Sciences*, *17*(1), 33–59.
Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology*, *80*, 1–14.
Mercer, N., & Fisher, E. (1992). How do teachers help children to learn? An analysis of teachers' interventions in computer-based activities. *Learning and Instruction*, *2*, 339–355.
Meyer, D. K., & Turner, J. C. (2002). Using instructional discourse analysis to study the scaffolding of student self-regulation. *Educational Psychologist*, *37*, 17–25.
Nathan, M. J., & Kim, S. (2009). Regulation of teacher elicitations in the mathematics classroom. *Cognition and Instruction*, *27*(2), 91–120.
Nathan, M. J., & Knuth, E. J. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction*, *21*(2), 175–207.
Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: University Press.
Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, *40*(1), 1–12.
Rasku-Puttonen, H., Eteläpelto, A., Arvaja, M., & Häkkinen, P. (2003). Is successful scaffolding an illusion? — Shifting patterns of responsibility and control in teacher–student interaction during a long-term learning project. *Instructional Science*, *31*, 377–393.
Scott, P., Mortimer, E., & Aguiar, O. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, *90*, 605–631.
Smit, J., & Van Eerde, H. A. A. (2011). A teacher's learning process in dual design research: Learning to scaffold language in a multilingual mathematics classroom. *ZDM, The International Journal on Mathematics Education*, *43*(6), 889–900.
Smit, J., Van Eerde, H. A. A., & Bakker, A. (in press). A conceptualisation of whole-class scaffolding. *British Educational Research Journal*. http://dx.doi.org/10.1002/berj.3007/pdf.
Smit, J., Van Eerde, H. A. A., Kuijpers, M., & Bakker, A. (submitted for publication). Development of a pedagogical genre and evaluation of students' genre proficiency: A linguistic turn in educational design?
Speer, N. M., & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education*, *40*(5), 530–562.
Van de Pol, J., Volman, M., Elbers, E., & Beishuizen, J. (in press). Measuring scaffolding in teacher – small-group interactions. In R. M. Gillies (Ed.), Pedagogy: New Developments in the Learning Sciences.
Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271–297.
Van de Pol, J., Volman, M., & Beishuizen, J. (2011). Patterns of contingent teaching in teacher–student interaction. *Learning and Instruction*, *21*(1), 46–57.
Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89–100.