Linguistics and Education 92 (2026) 101502

Contents lists available at ScienceDirect

Linguistics and Education

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

Using the body and material artefacts for spatial reasoning in classroom programming activities

Kreeta Niemi [a,*], Teppo Jakonen [b], Susanne Roos [c]

[a] University of Jyväskylä, 40014 Jyväskylä Finland
[b] University of Turku, 20014 Turku, Finland
[c] Teacher Training School, University of Jyväskylä, 40014 Jyväskylä, Finland

ARTICLE INFO

Keywords:
Programming
Classroom interaction
Conversation analysis
Embodied cognition
Spatial reasoning

ABSTRACT

We use multimodal conversation analysis to investigate programming activities as embodied and interactional phenomena. Drawing on the theory of embodied cognition, we analyse video-recorded interactions of Finnish pupils aged 9–10 years who are programming block-based codes for educational robots for a collaborative story animation project. The analysis focuses on how the pupils negotiate directional instructions, transform them into code, and implement their programming. The analysis reveals i) how the coding platform and its features constitute a situated problem and material space that the participants structure through embodied and social negotiation, ii) how the participants use their bodies as resources for cognitive reasoning and communicating the contingencies of the activity and iii) engage in object manipulation for testing and accommodating what is possible with the space and materials. The findings also indicate that a trial-and-error approach in which pupils test, observe and refine their code is essential to activity engagement and for developing programming and computational thinking skills.

1. Introduction

In today's classrooms, there is a growing push to cultivate computational thinking (CT) skills, which pupils usually practise in programming activities and, increasingly often, by using various kinds of hands-on educational robots (Bakala et al., 2021; Ben-Ari & Mondada, 2017; Chevalier et al., 2020). Programming a robot to perform an action involves creating and making sense of computational code, i.e., sequences of commands that direct the robot's actions and movements. Hands-on programming activities exemplify the material and embodied nature of learning to code: for instance, programming a robot to navigate a specific route inherently involves spatial reasoning regarding the robot's position, the intended direction, and distance, as well as taking into account features of the material environment and space that may prevent or afford the robot's movement. Thus, hands-on coding with robots can be understood as a process of converting spatial reasoning into precise instructions, i.e., code. In collaborative coding activities, this involves verbal and embodied reasoning between participants (see Berson et al., 2023).

This article explores spatial reasoning in coding activities as a form of interactional and embodied cognition (e.g., Shapiro and Stolz, 2019). We draw on multimodal conversation analysis (e.g., Mondada, 2019) to analyse video-recorded interactions of a pair of third-grade pupils as they program hands-on, block-based educational robots to move during a story animation activity. We examine how the pupils use their bodies and material objects as resources for spatial reasoning, for collaboratively producing computational code (e.g., go three blocks forward, turn right), and for instructing one another in spatial reasoning. Our aim is to advance educational research on CT by illustrating the interactional and embodied nature of programming in early learning contexts. In sequential analyses of selected data extracts, we also show how pupils' trial-and-error coding activities – the educational value of which previous research in robotics has questioned (e.g., Chevalier et al., 2020) – can in fact constitute a rich context for displaying and enacting embodied cognition for the purpose of understanding and viewing the material space and movement trajectories from the perspective of robots and their relation to each other and to space.

* Corresponding author.
E-mail addresses: kreeta.niemi@jyu.fi (K. Niemi), teppo.jakonen@utu.fi (T. Jakonen), susanne.roos@norssi.jyu.fi (S. Roos).

https://doi.org/10.1016/j.linged.2026.101502
Received 6 April 2025; Received in revised form 3 October 2025; Accepted 8 January 2026
Available online 6 February 2026

2. Computational thinking as an embodied, material and interactional phenomenon

The digital age and contemporary views on learning have foregrounded the importance of CT and programming skills in educational systems. Many Western countries have introduced programming into curricula with the aim of developing students' CT skills (Ben-Ari & Mondada, 2017; Berson et al., 2023). For example, in Finland, the context of the present study, the curriculum includes the notion of a 'digital compass,' which includes the constructs of understanding causalities, problem-solving, and operating and correcting instructions (Finnish National Agency for Education /Opetushallitus, 2022). At the same time, new coding tools have entered the educational technology market, including educational robots, which have been identified as promising for fostering students' CT skills and promoting their interest in computer sciences (Chevalier et al., 2020; Ham, 2018; Wing, 2006). Many coding robots, including the one investigated in this article, have been designed for early learners as easy-to-use and gamified tools that move in the spatio-material space (Critten et al., 2022). As an element of a technology-rich learning environment, hands-on coding robots can provide affordances for learners to use their bodily resources and experiences to promote understanding of abstract concepts in a concrete way (e.g., Niemi & Jakonen, 2022). Efficient use of the technology also requires teachers to have the skills needed to design and direct appropriate learning tasks with these tools.

Research on CT presents it as incorporating skills and knowledge related to various interconnected concepts and practices, such as problem decomposition, algorithmic thinking, pattern recognition, debugging, and testing (Bers et al., 2019; Chevalier et al., 2020; Wing, 2006). In programming activities, CT can be practised by creating a step-by-step instruction that outlines how a robot will perform each task. This methodical approach not only promotes a systematic understanding of programming concepts but also reinforces key principles of CT including determining its specific movements and actions (Ham, 2018). Instructing robots to move requires transforming directional instructions (e.g., turning right, left, advancing) into device-appropriate computational code that the robot can execute.

Creating this kind of directional and actionable computational code involves spatial reasoning, which refers to the ability to understand, visualise and manipulate objects and other elements and to understand how objects and elements relate to each other and appear from different perspectives (Berson et al., 2023). Spatial reasoning is translated into precise instructions for the robot, and then, as the robot executes the instructed action, students can observe what happens. If the desired result is not achieved, students can revise their instructions and try again.

Studies of CT have found this kind of trial-and-error approach to be typical in children's programming activities (Ben-Ari & Mondada, 2017; Chevalier et al., 2020), and it is also the most common approach in our data. Some researchers have questioned the educational value of trial and error, arguing that pupils who engage in such activities often spend an inordinate amount of time programming and evaluating without developing other necessary skills or problem-solving strategies. According to Chevalier et al. (2020) trial-error-approach may not effectively foster the development of critical skills, particularly in computational thinking. However, these studies usually treat programming as a series of discrete tasks and thus fail to capture the evolving interactional processes and collaboration between participants and their environment. This oversight can lead to a fragmented understanding of programming, whereas a process-oriented approach can allow more holistic understanding and the interplay of social, material and spatial elements.

Thus far, there is limited research that systematically takes into account the embodied and social nature of coding activities. One notable exception is, however, a study by Kopcha et al. (2021), which explored how young children's computational thinking is expressed through their embodied interactions with technology and others. The study showed that children used their bodies, gestures, and environmental structures to simulate robot movements and engage in algorithmic and multiplicative reasoning. In this study, we approach coding and especially spatial reasoning from a perspective that foregrounds their nature as embodied, interactional and social phenomena that are inseparably intertwined with the local material environment (Goodwin, 2007). This perspective finds resonance with various theorisations of embodied cognition and learning which emphasise the intertwined connection between cognition, body and environment and suggest that understanding and learning occur through physical experience and social interaction (e.g., Barsalou, 2008; Goodwin, 2007; Merleau-Ponty, 1962; Shapiro & Stolz, 2019). From this viewpoint, phenomena such as hand gestures, body movements, and the manipulation of material objects during coding activities can provide an entry point to examining embodied cognition from 'the participant's point of view' (i.e., the perspective of research participants, e.g., Ten Have, 2007). Much like social action, embodied cognition thus needs to be analysed with a sensitivity to how it relates to participants' situated concerns within an ongoing activity and how it emerges as they combine a broad range of resources (including talk and gaze) and affordances of the environment (Goodwin, 2003). The idea of embodied cognition is not new in the learning sciences (see Hall and Nemirovsky, 2012), but we aim to extend the perspective to the context of programming to shed new light on CT as embodied and social phenomena and especially to highlight the role of gestures.

3. Data and method

We analysed data selected from a larger research and digital skills development project examining schoolchildren's CT learning in interaction. A third-grade class of 22 Finnish 9–10-year old pupils and their teacher participated in the project for 12 weeks in the spring of 2024. The teacher planned the project, which was part of normal schoolwork, and this programming activity was included in the form of an integrated literacy project. Participation in the research was voluntary, and written consent for each participating pupil was obtained from their custodian (s) via information letters and consent slips. If a pupil had not wanted to take part in this study, they would have done the same activities without the data collection. The children worked in pairs, and each pair was videotaped with a separate GoPro video camera. The teacher aimed to form pupil pairs so that both pupils had similar literacy skills.

The children worked with a Matatalab classroom coding set, which an early learning coding material set in which codes are programmed with coding blocks (Fig. 1). The platform includes coding blocks with different kinds of functions, a control board, a nature map, a bot and a command tower. The bot is moved on the nature map by placing coding blocks on the coding board. When coding blocks are on the board, pressing the start button makes the command tower transmit the coded

Fig. 1. Matatalab coding set (reproduced from https://www.matatastudio.com/codingsetpro.html).

command to the robot via Bluetooth.

The coding activity was part of a literacy project in which the pupils had first written a story in pairs in their notebooks. The story had to have at least two characters going somewhere. The teacher told the pupils that they would later animate their story using Matatalab sets. The pupils were already familiar with the sets and how they are used, which likely helped them in writing and adapting their story for programming. They knew that in the animation, the characters in their stories would be represented by bots that would be made to move between the locations of the nature map. The story had to include dialogue between the characters and a plot with a climax, a resolution and an ending. After completing their story, the pupils created a mind map to visualise the bots' movements and paths in the story. Creating a visualisation involved breaking down their story into smaller parts, which in CT language is decomposing the story into code. After visualisation, the pupils started programming. The final stage of the project entailed the pupils themselves video-recording their animated story while reading aloud and programming at the same time. All sessions were video-recorded, creating a corpus of approximately 7 hours of video footage for each pair starting from the beginning of the literacy project.

We focus on one pair's programming activity in this article. We chose the pair because they spent almost all of their time working on the task at hand, with very little time spent engaging in off-task talk or having/resolving conflict. The selection made it possible to focus the analysis on coding interactions. Although both pupils in our focal pair participated actively in the programming, they also had a relatively stable division of labour so that one had the main responsibility for writing each of their code segments in the coding plan (used eventually for video shooting their final animation).

We analysed the pair's coding activities from the methodological perspective of multimodal conversation analysis (e.g., Mondada, 2019). It enabled us to examine their embodied cognition through the verbal, embodied and material interactional practices they employed to jointly negotiate directional codes for their bot in the situated context of the story animation. We initially viewed the focal pair's interactions through a data-driven process of 'unmotivated looking' to explore phenomena of interest for further examination (Ten Have, 2007). Our attention was drawn to how the pupils used their bodies and the available material artefacts to make sense of the bots' current orientations on the nature map, their directions, and locations they would be programmed to go to. Through repeated viewings of the video, we noticed recurring ways of using gesture and body orientation to accomplish coding, and we analysed these in a more systematic manner. We present our findings in section 4, using Jefferson's (2004) and Mondada's (2019) conventions to transcribe talk and embodied conduct. We use pseudonyms in the transcript and have treated transcript images with pencil sketch filters to maintain participant anonymity.

4. Results: Negotiating spatial routes for programmable robots

In this section, we analyse five representative extracts to demonstrate how the pupils created and negotiated computational code as they were programming bots to move on the coding platform. The focal pupils, A and B, are either sitting or standing side by side in the extracts, each using coding blocks and a command tower to operate their 'own' bot on a shared nature map in front of them on the table. The extracts illustrate different aspects of the interactional accomplishment of segments of code. We will show how the pupils proposed and planned the next 'codable' route for their bots (section 4.1), translated these negotiated directions and routes into segments of code that could be entered into the coding kit (section 4.2) and worked with written representations of code segments that they had previously recorded in their coding plan (section 4.3). Directing attention to how the participants used their bodies and the available material artefacts to make sense of spatial directions and instructions, we aim to demonstrate how coding in this kind of an activity context inherently involves embodied cognition.

4.1. The role of gestures in proposing and planning routes

Extract 1 illustrates the key role played by gestures in visualising and negotiating potential next directions and routes for the bots. Before the extract, B was away from the table for a short moment, during which A (left in Figure 1.1a) tested the combination of a so-called random movement block and the number block 5 on her bot. The random movement block (which the pupils referred to as the "run" block) makes the bot perform a surprise movement but return to the same location it started from. When B (right in Figure 1.1a) returns to the table, they test the combination simultaneously (lines 1–3). A then makes a proposal for the next route for the bots by using a combination of talk and gesture (lines 5–6). As the extract begins, the bots are side by side on the nature map, facing the same direction (as shown in Figure 1.1b).

```
01 B      okei, (.) än (0.3) [yy (.) tee (.) nyt.]#
02 A                         [yy (.) tee (.) nyt.]
          'Okay     N          O       W       now'
   fig                                               #fig1.1a/1.1b
```

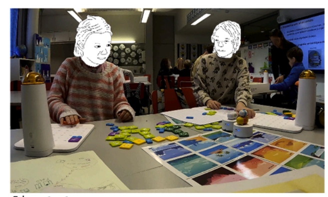

fig 1.1a

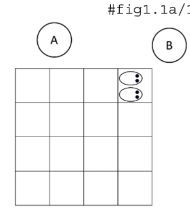

fig 1.1b

```
03        (3.2) ((A and B press start; the bots say 'bye bye')
04 B      [he he he he he he he he he he he he he he he    ]
05 A      [he he he he he he he .hh *joo hh   *sit ne menee]#=
                                    'Yeah     then they go'
   A                                *stands up*gestures->
   fig                                                          #fig1.2
06        ¤=kään#tyy tonne?#*=
           'turn there'
   A                        ->*,,,
   B      ¤gaze to her own board->
   fig          #fig1.3    #fig1.4
```

fig 1.2

fig 1.3

fig 1.4

```
07 B      =mitä* ¤tapahtuu jos laittaa juoksu nelja.
          'What   happens   if you put "run four"'
   A      ->,,,*
   B            ->¤puts a new block on her control board>>
08 A      nii määki mieti,
          'Yeah that's what I've been thinking'
09        ((A and B test "run four" combination))
```

Extract 1. Depicting a route with gesture.

As the participants are expressing their appreciation for the previous choreography of the bots through laughter (lines 4–5), A stands up and orients to moving on in the story. She takes a sharp inbreath, conveys her approval of the previous bot movements ('joo,' *yeah*) and proposes where the characters could 'then' go (line 5). The proposal combines a deictic expression ('tonne,' *to there*) and a gesture that identifies where *there* is by depicting the turning movement. A's gesture involves moving both hands from behind the two bots to above them so that the bots are under her palms and her hands are pointing to the same direction as the bots (Figure 1.2). In a single continuous movement, she then twists her hands towards the right (Figure 1.3) so that her palms are facing each other. The gesture is closely coordinated with her talk so that it ends in a brief gesture hold (as illustrated in Figure 1.4) just when A utters 'tonne' (*to there*). Altogether, the movement of A's hands thus provides a transient depiction of a right turn above the two bots, in a space that enables the recipient to visually examine the gesture by reference to the bots' positions and orientations on the nature map. Furthermore, using two hands for the gesture provides a visible representation of the two bots and the (in this case, aligned) trajectories that are being proposed for them.

In Extract 1, B does not respond to A's proposal. Notice that B shifts her gaze to her coding board on the table (line 6) and, because of this, does not visually orient to A's gesture. She then makes a competing proposal for testing the random movement coding block in combination with a different number block (line 7). Therefore, the lack of response by B does not appear to signal trouble understanding A's proposed route but is rather an indication of momentary misalignment due to an overlapping course of action. The gesturing in A's proposal provides thus both a resource for an action (proposal) addressed to B, even if she does not respond to it, and a resource for embodied cognition for the speaker herself, one that visualises a planned route. Extract 2 shows another instance of a gesture in proposing a route for a bot, this time in the form of 'finger drawing' the route on the nature map (lines 4–6). As in Extract 1, the brief duration of a transient gesture can involve trouble if there is a lack of mutual gaze during the gesture, and here B eventually makes the gesture visible to A through repair.

```
01          (3.2)             ¤(0.8)
    A       >>walks to desk->
    B       >>walks to desk-¤gaze to notebook->
02  B       öö* (.) mikä on sitten,¤
            'Erm    what's  then'
    A       ->*sits down, looks at blocks, picks up a character->
    B                             ->¤gz to nature map->
03          (0.4)
04  B       .hh ¤ja mennäänks me niinku *¤<↑tästä# näin,>¤
                 'And do we go  like      here like this'
    B         ->¤points at bot-----------¤finger draws---¤
    A                                ->*looks up, turns gaze to teacher->
    fig                                          #fig2.1a/b
05          (.)
06  B       ¤vai <näin.>*#
            'Or   like this'
    B       ¤finger draws->
    A                  ->*looks at the character in her hands->
    fig                        #fig2.2
```

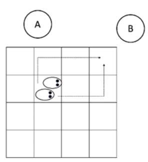

fig 2.1a fig 2.1b fig 2.2

```
07          (0.2)¤(0.4)
    B          ->¤
08  A       ¤häh?*
            'Huh?'
    B       ¤gaze to A->
    A          ->*lifts gaze, looks ahead->
09          (0.3)¤(0.1)
    B          ->¤gaze to nature map->
10  B       näin*¤kö,#
            'Like this'
    A         ->*gaze to bots->
    B          ->¤finger draws->
    fig                #fig2.3
11          (0.4)¤(0.6)
    B          ->¤
12  A       mm=
13  B       =¤vai #näinkö.¤
            'Or like this'
    B        ¤finger draws¤
    fig            #fig2.4
14          (0.2)*¤(0.6)
    A            *...->
    B             ¤...->
15  A         #*↑täl¤lein,  *¤
                'Like this'
    A        ->*finger draws*
    B             ->¤finger draws¤

    fig       #fig2.5
```

fig 2.3 fig 2.4 fig 2.5

```
16          (0.4)
17  B       eli kääntyy ((picks a coding block))
            'So (it) turns'
```

Extract 2. "Finger drawing" a possible route on the map.

Extract 2 begins as the participants return to the table with their coding plan. As B arrives at her place, she looks at the story notebook on the table during line 1 and orients verbally to the next codable bot movement in the story ('mikä on sitte,' *What's then*) in line 2. She then shifts her gaze to the nature map, on which the two bots are in the same cell, facing the same direction (Figure 2.1b). B requests A to confirm the route out of two alternatives (lines 4–6, indicated as dashed arrows in Figure 2.1b).

The alternative routes are presented through a multimodal action package that involves a combination of deictic expressions (*here, like this*) and a right-hand index finger gesture to 'draw' the two routes on the nature map (see Figures 2.1a and 2.2). This action can be seen as an example of embodied cognition, in that B uses bodily movement to carry out key cognitive operations: she explores possible paths, compares alternatives, and projects future robot behavior. The finger drawing does not merely accompany verbal reasoning but constitutes the reasoning itself, enabling B to simulate and evaluate potential movements from the robot's perspective within the material environment. However, as B is finger drawing the first route, A does not orient to it visually. Instead, A's gaze is on a nearby teacher during the depiction of the first alternative route (off-image in Figure 2.1a) and a story character that she has picked up and that she fidgets with during the proposal for the second route (visible in Figure 2.2).

A initiates repair (line 8) and subsequently shifts her gaze to the nature map (line 10). B re-does the alternative question in a starkly reduced form (lines 10, 13), accompanied by another instance of 'finger drawing.' This time, the drawn routes are visually available to A, whose gaze is directed at the nature map (Figures 2.3 and 2.4). A provides an early response after the first alternative in line 12 in the form of a non-lexical vocalisation (*mm*), which is somewhat ambiguous as it is a nonconforming response to an alternative question: It could be heard as either an acceptance of the first route proposal or a marker of receipt of B's turn. As the design of alternative questions resembles polar questions until the speaker produces the conjunction 'or,' such turns can pose problems of action ascription for participants (see Drake, 2021). Here, both participants orient to the ambiguity of the turn as B completes the second alternative: During the silence in line 14, they simultaneously move their hands to the bots (Figure 2.5) and use their index fingers to re-draw B's first alternative route, which preceded A's *mm*. A also provides a verbal confirmation (line 15), during which B withdraws her hand from the map and thus aborts the route she had begun to draw. B's subsequent verbalisation of the first movement of the bots (line 17, 'eli kääntyy,' *so turns*) and her proceeding to pick up a coding block shows that she treats the next route as having been decided. Indeed, as Figure 2.1b indicates, a left-hand turn is the first codable movement needed to proceed along the route that A finger draws in her response.

4.2. Transforming the next route into computational code

The previous section showed how the pupils made sense of and conveyed directions and locations on the nature map by using gestures that are carefully timed with talk. Besides offering resources for planning routes for the bots, the body is a central resource for formulating segments of code for moving the bots on the nature map. To move the bots, the participants need to establish an equivalence between directions (such as the dashed arrows in Figure 2.1b of Extract 2) and the arrow blocks (forward, left, and right) of the coding set that make a bot move when placed on the coding board. This is a task that requires spatial reasoning and cognition as it requires the programmer to attend to the bot's orientation on the nature map and how many steps/cells it ought to move in the desired direction.

This section shows how the pupils use their bodies, the surrounding space, and the material resources available to them to design code segments for their animation. Extract 3 illustrates a situation in which the bot's and the participant's points of view are aligned. As the extract begins, the robots are side by side, facing away from the pupils (as shown in Figure 3.1b). In such a constellation, the visual perspectives of the pupils and the bots are aligned: What is 'left' and 'right' is the same for the pupils and the bots. Before the extract, the pupils have agreed to try to make the bots talk to each other and have already placed the “random movement” coding block on their coding boards. Following a long negotiation, the extract begins as A agrees to B’s earlier proposition (line 1) to make the bots first turn towards each other before performing the ‘random movement.’ This movement trajectory is precisely what A begins to decipher in the extract with the help of gestures.

```
01 A     no, (.) laitetaan sitte¤     (.)     *kääntyy¤
         'Well   let's then make (them)        turn'
   A     >>repositions block on control board*
   B     >>erases text---------¤...................¤writes->
02 B     käännös oikeel*le,#
         'Turn to right'
   A                      *lh points forward, palm downwards->
   fig                         #fig3.1a/b
```

fig3.1a

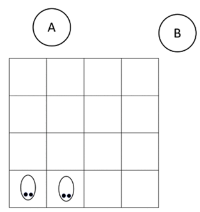

fig 3.1b

```
03       (0.9)
04 A       *sä #käännyt-*
            'You turn-'
   A     ->*lh palm up--*turns lh wrist to point left->
   fig             #fig3.2
05       (0.2)^   #(3.8)            ^(0.3)*(0.2)
   A          ^takes a block with rh,
               puts it on her board^
   A                                   ->*turns twd B->
   fig               #fig3.3
```

fig 3.2

fig 3.3

```
06 A     ->*↑sä et #käänny- ^sä käännyt *<vasemma>-# ¤eiku* oikeelle#
             'You don't turn- you turn lef- no right'
   A     ->*lh points forward-----------*twists lh left---*takes a block
                                                            from table->
   A                         ^rh points forward, then right->
   B                                              ->¤stops writing,
                                                      gz-bot->
   fig               #fig3.4                      #fig3.5         #fig3.6
```

fig 3.4

fig 3.5

fig 3.6

```
07 B     °(ö)°¤*
          'Err'
   B       ->¤gz to text, writes->
   A        ->*glances at B's board, then the block->
08       (0.6)*(1.5)
   A       ->*puts block on B's board->
09 A     ↑mä käännyn <vasem°paa,°>
         'I turn left'
10       (0.2)¤(0.4)#(3.6)¤*(4.8)¤   (0.8)     ¤(1.0)*(0.7)         ¤(0.5)
   B       ->¤gz-ahead,
                gestures¤writes¤gz-board,
                                 stops writing¤glances at bots,
                                              hand on start button¤gz-A>>
   A                ->*                            *hand on
                                                    start button>>
   fig              #fig3.7
```

fig 3.7

```
11 A     än, (.) [yy tee (.) nyt] ((pushes button))
         'N       O  W       now'
12 B             [yy tee (.) nyt] ((pushes button))
                 'O  W       now'
```

Extract 3. Using gestures to design directional code.

As A agrees to coding the bots to turn to each other (line 1), she is simultaneously repositioning the random movement block on her board to make space for a new (turning direction) block at the beginning of the coding segment. She thus orients to revising her code. B also treats the turning of the bots as a decided matter, as is evident from how she erases text from the coding plan during A’s turn and begins to revise it. B also utters a ‘codable’ command in line 2 (‘käännös oikeelle,’ *a turn to the right*), which is potentially what she also writes down in the plan.

A begins to establish the directions in which the bots should be coded to turn so that they will face each other. During the silence in line 3, she looks at the bots and points forward with her left hand, indicating the direction that represents straight ahead for both bots (Figure 3.1a). As A begins a turn that projects an instruction for B (‘sä käännyt,’ *you turn*, line 4), she first turns her left palm upwards (Figure 3.2) and then twists her wrist so that the hand forms an angle that points to the left (Figure 3.3). The gesture serves as an embodied completion of her spoken turn (Olsher, 2004) and is hearable as an attempt to formulate the code for B. However, during the silence in line 5, A also picks up a left arrow coding block from the table and puts it on her coding board. Given that the left-turn arrow corresponds to A’s angled hand gesture, it seems that she is employing the gesture as a heuristic tool for reasoning her *own* next code, not B’s code.

Having placed the coding block on her own board, A turns to B and re-does her instruction for the next movement of B’s bot in line 6, through multiple cutoffs and restarts. While talking, she first points forward with her left hand (Figure 3.4) and then, in a single trajectory, also moves her right hand forward and then immediately to the right so that her right hand makes an angular shape mid-air. As she is turning her right hand, her left hand turns slightly to the left, leaving her hands in a position in which they point in separate directions (Figure 3.5). Simultaneously, she says *you turn lef*– (‘sä käännyt vasemma–’) but quickly repairs it to *right* (‘oikeelle,’ line 6; see Figure3.6). Maintaining her right hand in the angled position, she then uses her left hand to pick a coding block from the table and puts it on B’s coding board. For most of the extract, B’s orientation is on the coding plan in which she writes. However, in line 6, she pauses writing and quickly glances at the bots. Some time later in line 10, she pauses writing for four seconds, looks ahead and holds her hands in mid-air pointing forward (Figure 3.7). The gesture can be seen as a representation of direction, and the fact that B then resumes writing without saying anything suggests that she uses the gesture as a resource for her code writing. Towards the end of the long silence in line 10, the students prepare to implement the code by moving their hands on the start button of their programming boards. Through the practice of a shared countdown (lines 11–12), A and B ensure that they press the start button at the same time in order to make the bots move synchronously.

In sum, Extract 3 shows how the ‘arrow’ gesture can serve as a visualisation tool for designing segments of code. In a spatial configuration in which the participant’s body orientation is aligned with that of the bot (and the control board), such a gesture can be a simple yet effective tool for establishing what kind of arrow block is needed next. However, depending on the material arrangements of bodies, bots, and the coding boards, visualisation may involve more complex embodied practices. Extract 4 illustrates this in a situation where the orientation of the robot differs from that of the students. As the extract begins, A and B have just watched their bots complete a movement trajectory that they previously coded. When that movement ends, A’s bot is positioned on the nature map so that it is facing A (see Figure 4.1b). In the extract, both participants are focused on designing the next code to be programmed for A’s bot to move on the board.

```
01 B      ↑just,
          'Exactly'
02 A      ↑toi o hyvä hh
          'That's good'
03        ¤(0.3)
   B      ¤takes a coding block, puts it on her control board->
04 A      .hh sitte- *oota#- (.) älä-¤
          'Then-      wait (.) don't-'
   A                 *raises hands in mid-air->
   B                                     ->¤
   fig                            #fig4.1a/b
```

fig 4.1a

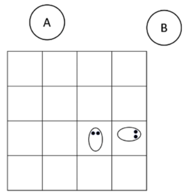

fig 4.1b

```
05          *(0.6)#
   A      ->*moves rh forward, stands up->
   fig            #fig4.2
06 A      mihin *suuntaa,#
          'In which direction?'
   A          ->*turns body away from the board, rh points->
   fig                   #fig4.3
07        (0.4)*
   A         ->*angles rh wrist->
08 B?     °mm°
09        (0.6)#
   fig        #fig4.4
10 B      ¤sä *käännyt# öö, (1.8) ¤oikee[lle.]
          'You turn err            to the right'
11 A                                   #*[ tän]ne (.) tämä hh
                                          'Here       this'
   B      ¤rh pinching gesture,
          moves hand to right-----¤
   A        ->*takes a block with lh,
          brings it next to rh---------*lifts block, puts it on board->
   fig                #fig4.5a/b           #fig4.6
```

fig 4.2

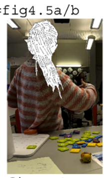

fig 4.3

fig 4.4

fig 4.5a

fig 4.6

fig 4.5b

```
12        (0.3)*(0.7)
   A         ->*
13 B      än yy tee n::::yt ((A and B press control board start buttons))
          'N  O W    NOW'
14        (1.4) ((bots move))
```

Extract 4. Reconfiguring body orientation to perceive directions from the bot's perspective.

The pupils begin to transition to the next coding cycle after assessing (in an approving manner) their previously coded bot movements (lines 1–2). B places a new coding block on her control board (line 3), and A orients verbally to a transition ('sitte,' *then*) in line 4. However, A quickly cuts off her turn and urges B to wait, lifting her hands mid-air as if displaying gesturally a need to suspend the activity (see Figure 4.1a).

A begins to form the code for her bot during the silence in line 5 by using her body as a resource. She stands up and first points with her right hand towards her bot on the nature map (Figure 4.2) and, in a continuous movement, twists her upper body away from the table (Figure 4.3). In that position her hand now points away from the bots to the same direction that her bot is facing on the nature map. Along with the body movement, she verbally problematises the next direction to be coded through an interrogative question (line 6).

Facing away from the bots and the nature map, A turns her right wrist so that her arm and hand form an approximately right angle (90 degrees) and her fingers point to her right (Figure 4.4). This combination of torqued upper body and wrist enables A to accomplish two things: to align her body orientation with that of her bot on the nature map and to visualise the next programmable direction (right turn) with the help of her arm shape. Holding her body and the gesture, A then employs her left hand to pick up a right arrow coding block from the table (Figure 4.5) and brings it next to her right hand (Figure 4.6). Gazing at the block and the pointing hand next to each other enables a visual comparison of the direction of the arrow to the shape of her pointing hand. In line 11, A's demonstrative verbal tokens ('tänne, tämä,' *here, this*) display that she treats them equivalent. She quickly "taps" the block in the air and places it on her coding board.

B's conduct shows that she treats A's turn (line 6, 'mihin suuntaa,' *in which direction*) as a request for help. In line 10, using the singular second person pronoun ('sä,' *you*), B tells A the direction to turn to. In producing the verbal instruction, B also uses gesture as a resource for spatial reasoning but makes a very different kind of gesture. Unlike A, B does not turn away from the map but instead makes a pinching gesture in front of her (see Figure 4.7) and turns her body towards the right, thus embodying a movement to the right. The fact that B verbally completes her instruction to A only after she has completed her gesture during the 1.8-second-long turn-internal silence suggests that the gesture serves as a resource for her spatial reasoning. Unlike gestures that are designed to display understanding to the recipient (see e.g., Jokipohja, 2023) here B does not appear to visually monitor how A receives the gesture, which suggests it is primarily self-oriented.

All in all, Extract 4 illustrates how the use of embodied and material resources enable the participants to design and visualise directions to be coded for the bots. In this activity and with these material resources, the coding of a desired route requires spatial reasoning to establish a correspondence between the orientations of the bot on the nature map, one's own bodily orientation, and the directions pointed at by the arrows on the coding blocks. In other words, one needs to 'see' the nature map from the perspective of the bot to decipher where, for example, a right-turn arrow coding block would take it next on the map. Extracts 3, 4 demonstrate that for this kind of spatial reasoning, the body provides an available and easily malleable resource.

4.3. Implementing previously designed code

In discussing our final example, we focus on the role of the written coding plan – which the participants have been designing and writing in Extracts 1–4 – as an accountable structure of the animated story. Extract 5 shows a situation towards the end of the coding session when A and B have designed the code for all movements included in their short story. They are rehearsing for eventually video-recording the animation so that they implement the coded sequences one at a time and accompany the ensuing bot movements by reading aloud the relevant passage from the story. The bot movements immediately preceding Extract 5 have been problematic because the pupils have accidentally fixed B's story character on top of the bot slightly diagonally. The pupils have used the orientation of the character to establish what is straight ahead, left or right for the bot but have not noticed its difference from the bot's orientation, which has eventually resulted in the bot going diagonally across the cells in the nature map (see Figure 5.1a). A has already expressed some frustration and bewilderment at the situation, and in line 4 she blames the troubles on the pair's coding plan, which B has largely written.

```
01          (3.7) ((robots move asynchronously; A's robot goes off target))
02 A        no ei tää nyt mee yhtää hyvi,
            'This is not going well at all'
03          (2.8) ((A corrects robots' position, empties her coding board))
04 A        *↑MItä# sitten *ku emmä saa tästä ¤selvää.#
            'What then      cos I can't make sense of this'
   A        *gaze to B-----*turns to coding plan, picks it up->
   B                                               ¤...->
   fig              #fig5.1a/b                          #fig 5.2
```

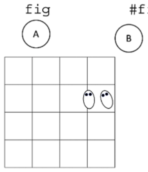

fig 5.1a

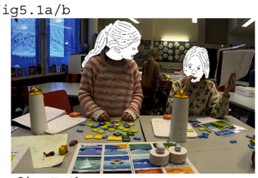

fig 5.1b

fig 5.2

```
05          (0.7)*¤
   A           ->*
   B         ...->¤takes coding plan from A->
06 B        no em↑määkää saa tästä ¤°selvää yhtää°.
            'Well I can't make any sense of this either'
   B                            ->¤scrutinises coding plan->
07 A        niinno *miks sää# (oot) sit kirjottanu.*
            'Well yeah why have you then written'
   A              *gaze to robots-----------------*turns around, lh-points->
   fig                      #fig 5.3
08          (0.9)
09 B        koska- *(0.9) #no ↑miks ¤ite oot kirjottanu °( )°=
            'Because-     Well why have you yourself written ( )'
   A           ->*turns lh wrist, picks up a 'left turn' coding block->
   B                             ->¤arranges coding plan and story in lh->
   fig                    #fig 5.4
10 A        =EMMÄÄ oo sitä kirjottanu.
            'I haven't written that'
11 B          *no ↑kyl sää oot tän* kirjot*tanu.
            'Well yes you have written this'
   A        ->*puts block on board*       *looks at coding plan->
12          (0.4)
13 A        no nii *s:en# (.) mutta* en tätä.*
            'Well yeah that but not this'
   A           ->*points at plan *          *moves hand above blocks ->
   fig                  #fig 5.5
14          (2.1)
15 B          *(sitten) kääntyy vasemmalle,* (.) ¤suoraan* ↓kolmone.#
            'Then turns to the left              straight ahead three'
   A        ->*takes left turn, swaps it for
                        another on B's board*           *rh on start button->
   B                                            ¤gaze to coding board->
   fig                                                             #fig5.6
```

fig 5.3

fig 5.4

fig 5.5

```
16          (0.5)*(0.3)
   A           ->*picks a 'straight ahead' block, puts it on her board->
17 A        ai nii joo (.) ¤hups.
            'Oh yes. Oops.'
   B                      ->¤gaze to A's board->
18          (0.5)*(1.5)*                               *¤(0.9)
   A           ->*picks a 'straight ahead' block,
                          puts it on B's board---*...->
   B                                                ¤gaze to own board->
19 A        *°°yks¤ [kaks    #]kol.°°*¤
20 B                [°°°( )°°°]
               'One two three'
   A        *points at map-----------*picks a '3' block, puts on own board->
   B           ->¤glances at plan----¤picks a '3' block, puts on own board->
   fig                     #fig5.7
```

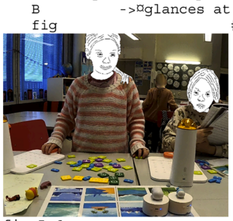

fig 5.6

fig 5.7

```
21          (0.9)
22 B        (mulla/sulla) on jo¤ kolmone.=
            'I/you already have three'
   B                         ->¤
23 A        =joo¤joo.*
            'Yeah yeah'
   A              ->*rh on start button, gaze to B's board->
   B          ->¤rh on start button, gaze to A's board->
24          ((A and B press the start button))
```

Extract 5. Topicalising the written coding plan.

A topicalises the written coding plan as part of a prosodically emphatic display of frustration (line 4). With a turn-initial gaze shift, A addresses her question to B (Figure 5.1b) but immediately shifts her gaze to the coding plan on her other side, picks it up and begins to investigate as the turn is in production (Figure 5.2). The loudness and high pitch at the beginning of the turn contribute to framing the question as insisting. A's gaze shift to the plan is an observable display of 'doing scrutinising,' which she performs simultaneously as she complains that she cannot make sense of the text. Examining the plan is therefore an essential component of how A holds B accountable for the complaint. It demonstrates to B that A has made an effort, albeit briefly and possibly late in the complaint turn, to comprehend the text.

A's treatment of the plan as unclear creates an implication of blame. B reaches out to the plan (visible in Figure 5.2), takes it from A, and reciprocates and upgrades ('selvää yhtää,' *any sense*) A's claim. Similarly to A, she begins the verbal claim before she observably scrutinises the written text. This leads to a kind of a tit-for-tat exchange in which the participants accuse the other of having written the unclear text (lines 7–13). As the dispute is ongoing, A begins to design the next code without the written coding plan: she peeks at the robots' orientation as B is scrutinising the coding plan (Figure 5.3), aligns her body with the robots' orientation like in Extract 4 and twists her left wrist to form a left-turn coding arrow shape (Figure 5.4). She picks a corresponding left-turn arrow block and places it on her board. In between placing these blocks on her board, she orients to the coding plan briefly as she counters (line 13) B's accusation that she has written the unclear code (line 11). As part of that, she identifies by pointing (Figure 5.5) which part of the text B has written (and the unclarity of which is therefore B's responsibility).

B reads aloud the pair's previously written code in line 15. At this point, A has already placed a block on her own board and one on B's board and has her hand above the start button (Figure 5.6). This suggests that she is getting ready to implement the code sequence to move the bots. However, B's reading the written code aloud occasions a correction of A's programming, which has so far only included a left turn. A adds a straight-ahead arrow on her own board (line 16) and uses the particle chain 'ai nii joo' (*oh yeah*) to receipt B's read-aloud as a known state of affairs, as if conveying that she now remembers the coded sequence when she is being reminded of it (line 17). She adds a corresponding straight-ahead arrow block on B's board and almost silently counts the distance that the bots should travel (line 19), briefly tapping her index finger in air three times (Figure 5.7) in synchrony with the uttered numbers. Given that B has already told A the distance from the coding plan, counting it again seems to treat the coding plan as something less than the primary instructional document in this situation.

5. Discussion

In this article, we have used multimodal conversation analysis to investigate interaction during hands-on, block-based coding activities as part of a broader literacy project with a third-grade class. Our analysis has focused on how the pupils use their bodies, the classroom space and the material objects of the coding kit as resources for the spatial reasoning needed to program their robots to move from one place to another. The analysis has shown how the pupils made sense of spatial directions as they planned directions and routes for the robot to traverse (section 4.1), transformed these into block-based computational code (section 4.2), and implemented previously recorded written representations of the code (section 4.3).

The extracts illustrate the interactional and embodied nature of programming the bots to move on the board along a desired path. Programming them required understanding and viewing the material space and imagined movement trajectories from the perspective of the bots and analysing the bots' positions in relation to each other, to the pupils themselves and to features of the material environment. The pupils manipulated objects in relation to their own bodies and the environment of the nature map and the material space around them. The material environment thus not only served as a situated resource for interaction and achievement for the coding task but also constituted its immediate context and a practical problem (see also Greer et al., this issue). These pupils found structure in this problem largely through an embodied manner co-occurring with talk.

Embodied displays of cognition was evident in all the extracts analysed. The pupils' use of their fingers and hands in engaging bodily with the task materials is noticeable and integral to the cognitive processes involved in problem-solving and situational understanding. The pupils' hands worked as heuristic tools to explore and manipulate tangible coding tools, helping them reason and depict directions and routes. Sometimes, this might take the form of using fingers to draw possible routes in the air (Extract 2), pointing in directions (all extracts), or counting the distance a bot had to travel by tapping their fingers in the air in sync with the talk (Extract 5). As CA studies have demonstrated, gestures are motivated by, and linked to, the contingencies of the task at hand and the material environment, designed to be publicly visible and interpretable by the recipient in task-relevant ways (Goodwin, 2003; Lilja & Jokipohja, 2024).

In our coding context, gestures were an important means of enacting and displaying cognition for negotiating and solving task-related problems. This embodied reasoning underscores the need to recognise embodied learning, particularly in disciplines like coding and CT, where challenging, abstract concepts can be playfully grasped. Regarding the role and timing of embodied resources in spatial reasoning, the analysed extracts show that, in particular, A's process of establishing codable directions was often initiated with embodied conduct prior to verbal expression (Extracts 2–4). Directions and movements in space were negotiated at first bodily, then verbally, and then transformed into block-based coding language with coding blocks. As well as representing the code on the coding platform, coding blocks helped in reasoning and building the code.

The verbalised code served as an instruction for the other participant, facilitating understanding and implementation of tasks, and as a cognitive tool for reasoning and social interaction. Even when verbalisation of the code was more akin to self-talk, it also served as a resource available to the other although it was not specifically addressed to them. Although the language was reminiscent of giving instructions ('Turn right'), the instructions were still open to negotiation, indicating a distribution of agency. The code itself was not simply a vehicle for conveying information on directions but embodied a local understanding of the features of the material and spatial environment and the written story framing the activity.

All in all, our findings challenge the view of trial and error as detrimental to CT (e.g., Chevalier et al., 2020) and suggest that trial-and-error practices should not be viewed merely as mechanical and repetitive actions, or loops, that yield minimal advancement in learning strategies. Instead, they constitute rich sites for social and embodied cognition as pupils interactionally co-construct and coordinate coding activities. When pupils test, observe and refine their code, they deploy social and embodied cognition as they negotiate with each other and use material artefacts collaboratively, which should be seen as an interactive achievement rather than a deficit model of programming activity. The observations in this study suggest that trial-and-error practices are an essential part of learning programming and developing CT skills (cf. Antle et al., 2013: Chevalier et al., 2020: Shute et al., 2017). When pupils can reason, monitor, observe and learn from their decisions, this can foster self-efficacy and ownership of their learning processes. In a large part of our dataset, pupils' trial-and-error 'loops' appear as a motivating and engaging form of coding, which is visible in mutual laughter and other expressions of appreciation (see e.g., Extract 1).

Furthermore, the trial-and-error practices observed in the data can be seen as instances of pupils' creative thinking. The coding activity was an integrated classroom literacy project. The story written beforehand served as a 'glue' which held the activity together, providing an instructional framework and meaning for the activity. The interplay between writing and coding highlighted the importance of storytelling in early learning coding activities and seemed to enrich the learning experience. On the other hand, the pupils continuously adjusted the story while programming on the basis of what they deemed possible to code with the platform. The orientation in the project can thus be seen as an example of a pedagogy of multiliteracies in which language use and meaning-making are embedded in and emerge from engagement with (digital) media and tools (e.g., Zapata et al., 2023). The carefully planned learning task scaffolded the learning activity and did not require much help from the teacher. Hands-on activities catered to different learning styles and thereby met the needs of today's classrooms. Altogether, the broader activity context of coding appeared to be very engaging to pupils beyond the pair we have discussed, as it also motivated exploration of the range of possibilities in the coding set. Future research should continue to explore the intersections of embodied cognition, spatial awareness, materials and collaborative practices in programming education in order to pave the way for pedagogical approaches that reflect the complexities of learning in an increasingly digital era.

CRediT authorship contribution statement

Kreeta Niemi: Writing – original draft. **Teppo Jakonen:** Writing – original draft. **Susanne Roos:** Writing – original draft.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kreeta Niemi reports financial support was provided by Research Council of Finland. Teppo Jakonen reports financial support was provided by Research Council of Finland. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Research Council of Finland grant numbers [356181] and [343480].

Data availability

The data that has been used is confidential.

References

Antle, A. N. (2013). Exploring how children use their hands to think: An embodied interactional analysis. *Behaviour & Information Technology, 32*(9), 938–954.

Bakala, E., Gerosa, A., Hourcade, J. P., & Tejera, G. (2021). Preschool children, robots, and computational thinking: A systematic review. *International Journal of Child-Computer Interaction, 29*, Article 100337.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645.

Ben-Ari, M., & Mondada, F. (2017). *Elements of robotics*. Springer Nature.

Bers, M., González-González, C., & Armas-Torres, M. B. (2019). Coding as a playground: Promoting positive learning experiences in childhood classrooms. *Computers & Education, 138*, 130–145.

Berson, I. R., Berson, M. J., McKinnon, C., Aradhya, D., Alyaeesh, M., Luo, W., & Shapiro, B. R. (2023). An exploration of robot programming as a foundation for spatial reasoning and computational thinking in preschoolers' guided play. In *Early Childhood Research Quarterly, 65* pp. 57–67).

Chevalier, M., Giang, C., Piatti, A., & Mondada, F. (2020). Fostering computational thinking through educational robotics: A model for creative computational problem solving. *International Journal of STEM Education, 7*, Article 39.

Critten, V., Hagon, H., & Messer, D. (2022). Can pre-school children learn programming and coding through guided play activities? A case study in computational thinking. *Early Childhood Education Journal, 50*(6), 969–981.

Drake, V. (2021). Alternative questions and their responses in English interaction. *Pragmatics, 31*(1), 62–86.

Finnish National Agency for Education /Opetushallitus. (2022). *Digital competence descriptions / Digitaalisen osaamisen kuvaukset*. Opetushallitus.
Goodwin, C. (2003). The semiotic body in its environment. In Coupland, J., & Gwyn, R. (Eds.), Discourses of the body (pp. 19–42). Palgrave/Macmillan.
Goodwin, M. H. (2007). Participation and embodied action in preadolescent girls' assessment activity. *Research on Language and Social Interaction, 40*(4), 353–375.
Greer, T., Bradn, A. & Nanbu, Z. (this issue): Giving Directions in Virtual Reality: Resources for Fragmented perceptions.
Hall, R., & Nemirovsky, R. (2012). Introduction to the special issue: Modalities of body engagement in mathematical activity and learning. *Journal of the Learning Sciences, 21*(2), 207–215.
Ham, D. A. (2018). Spatial thinking as a path towards computational thinking. In H. Ozci-nar, G. Wong, & H. T. Ozturk (Eds.), *Teaching computational thinking in primary education* (pp. 103–122). IGI Global.
Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). John Benjamins.
Jokipohja, A. K. (2023). Demonstrating and checking understanding—Bodily-visual resources in action formation and ascription. *Journal of Pragmatics, 210*, 87–108.
Kopcha, T. J., Ocak, C., & Qian, Y. (2021). Analyzing children's computational thinking through embodied interaction with technology: A multimodal perspective. *Educational Technology Research and Development, 69*(4), 1987–2012.
Lilja, N., & Jokipohja, A. K. (2024). Gestural depictions in requests for objects. *Language & Communication, 99*, 159–173.
Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Routledge.
Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics, 145*, 47–62.
Niemi, K. & Jakonen, T. (2022). Pelillisyys ja oppimateriaalit vuorovaikutuksessa. Bee-bot-robotti alkeisohjelmoinnin moraalisena artefaktina. /Gamification and learning materials in interaction. The Bee-Bot robot as a moral artifact in early programming education. In Heinonen, P. & Karvonen, U. (Eds.): Katso kuule koulua. Kehollisuus, materiaalisuus ja tekstit koulun vuorovaikutuksessa. / See, Hear the School: Embodiment, Materiality and Texts in School Interaction (pp. 93-115). Vastapaino.
Olsher, D. (2004). Talk and gesture: The embodied completion of sequential actions in spoken interaction. *Second Language Conversations/Continuum*.
Shapiro, L., & Stolz, S. A. (2019). Embodied cognition and its significance for education. *Theory and Research in Education, 17*(1), 19–39.
Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computationalthinking. *Educational Research Review, 22*, 142–158.
ten Have, P. (2007). *Doing conversation analysis*. Sage.
Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35.
Zapata, G. C., Kalantzis, M., & Cope, B. (Eds.). (2023). *Multiliteracies in international educational contexts: Towards education justice*. Taylor & Francis.