Teacher Orchestration of Language and Gesture in Explaining Science Concepts in Images

Thu Ngo [1] · Len Unsworth [1] · Michele Herrington [1]

Abstract

Students' difficulties interpreting diagrams remain a concern in science education. Research about improving diagram comprehension has included few studies of teachers' orchestration of language and gesture in explaining diagrams—and very few in senior high schools. Research with younger students and studies of research scientists' practice indicate the significance of the interaction of teachers' gesture and language in explaining visualisations. The strategic deployment of such teacher-focussed authoritative explanations has been observed in facilitating progression to more complex and symbolic representations in classroom work. However, the paucity of such research in senior high school leaves open the question of how these teachers use gesture and language in managing the challenges of explaining the intricate sub-microscopic and abstract visualisations senior high school students need to negotiate. In this paper, we outline existing studies of teachers' use of gesture and language to explain complex images in senior high school and investigate how it is managed by two biology teachers with images of different types and complexity representing the activity of certain cell components in the early phase of cell duplication. Implications are drawn for foci of further research including the role of a metalanguage describing different types of visualisations and their affordances.

Keywords Gesture · Multiple representations · Science explanation · Disciplinary discourse · Multimodality

✉ Len Unsworth
len.unsworth@acu.edu.au

Thu Ngo
Thu.Ngo@acu.edu.au

Michele Herrington
Michele.Herrington@acu.edu.au

[1] Australian Catholic University, Sydney, Australia

Introduction

The difficulties experienced by students in interpreting diagrams remain a prominent focus in science education research (Brandstetter et al., 2017; Kottmeyer et al., 2020; Kragten et al., 2015; Ruiz-Gallardo et al., 2019). Such difficulties are more common among educationally disadvantaged students from low socio-economic status (SES) backgrounds and/or whose first language is not the language of instruction (Bergey et al., 2015; Cromley et al., 2013; Nunes et al., 2017; Thomson et al., 2017). Working with diagrams was identified as problematic for students in our study of multimodal literacy development to address educational disadvantage in senior high schools in low SES areas (Unsworth et al., 2021). Among studies concerned with improving students' diagram comprehension, little attention has been given to how teachers use gesture and language to explicate diagrams in classroom contexts, especially in senior high school. However, the importance of apprenticeship in learning to read complex images via explication by a more experienced viewer mediating the visual representation through language and gesture has been emphasised by Roth and Pozzer-Ardenghi (2013). They report this as routine in biology research laboratories, detailing, for example, how an experienced research team member explained to newly recruited researchers how to read complex images of the project data. They emphasise how this involved pointing to, naming and tracing along the key components, thus highlighting gesture as integral to the coordination of language and image in this communicative interaction. From their video data of a senior high school biology class, Pozzer-Ardenghi and Roth (2007) argue that the teacher's coordination of language, gesture and image is teaching students how to look at and interpret the two-dimensional visualisation in relation the three-dimensional phenomenon represented. Moreover, the teachers' gestures are connecting the technical language being introduced with the visual representations of the entities, activities and properties of the phenomenon. Roth and Pozzer-Ardenghi (2013) propose that this orchestration of gesture, language and image is 'the very source for perceiving images in didactically useful and appropriate ways' (p.50). In this paper, we outline existing studies of teachers' orchestration of language and gesture in explicating complex images in senior high school classrooms. We then present our analysis of brief video excerpts of two teachers explaining different types of images representing similar phenomena in senior high school biology, from which we draw attention to key foci for further such research and for teacher professional learning.

Related Research

Extensive research advocating the use of multiple representations has tended to privilege the role of language in combination with various forms of visualisations and models but has also acknowledged the importance of gesture (Gilbert & Treagust, 2009; Tang & Danielsson, 2018; Treagust et al., 2017; Treagust & Tsui, 2013). The vast majority of these studies focussed on students' use of gesture in producing multimodal representations (Givry & Roth, 2006; Herrera & Riggs, 2013; Mathayas et al., 2019; Pantidos et al., 2017; Plummer et al., 2016; Siry & Gorges, 2019; Ünsal et al., 2018; Yeo & Gilbert, 2017). Early studies of teachers' gestures by Kress et al. (2001) in junior high school classrooms and Pozzer-Ardenghi and Roth (2007) in senior high school biology drew attention to meanings revealed through gesture that are not communicated visually or verbally, emphasising however, that meaning emerges from a process of interaction, contrast and conjunction of speech, gesture, image and material

resources such as models. These researchers argue that this interweaving is largely achieved through teachers' deictic gestures of 'pointing' and their use of pronominal reference and demonstrative pronouns in language, because they establish linkages between the teacher's talk, physical artefacts, visualisations and other models. Hence, gesture is a resource that can assist in building students' meta-semiotic awareness of how meaning is constructed through coordination of the distinct affordances of different types of representations.

Very few studies have addressed teachers' coordination of language, gesture and image. Padalkar and Ramadas (2011) investigated a teacher's use of researcher-designed gestures illustrating spatial concepts related to teaching the Sun-Moon-Earth system to year 8 students. Pre-designed gestures and their use with models and images were outlined, but there was no analysis of the teacher's moment to moment deployment of the gestures. Detailed investigations of language and gesture used with teacher-generated diagrams by Márquez et al. (2006) explaining the water cycle to year 7 students, and by Danielsson (2016) explaining dynamic and static representations of the atom in high school chemistry, focussed on the nature and extent of the contribution of each mode to the meaning of the science concepts and the convergence or divergence of the meanings realised by the affordances of each mode. Without attention to deictic gesture and verbal pronominal reference to track the identity of the scientific entities across the different modes, the approach in these studies obviated examination of how the teachers coordinated the multimodal meaning-making.

Within the little research on teachers' gesture, very few studies occurred in senior high schools. In their analysis of a year 12 biology teacher's coordination of a 3D model, diagram, language and gesture in explaining the anatomy of the heart and the role of its components in the circulatory system, Pozzer-Ardenghi and Roth (2007) showed how meaning is made in and through the inter-relationship of the multiple representations (p.103). This study focussed on the representation of macroscopic phenomena, and, apart from Danielsson's (2016) analysis of teachers' explications of their drawings of the atom, we found no other studies involving sub-microscopic visualisations or teachers' comparative explication of different diagrams of the same phenomenon. While the importance of teachers' orchestration of language and gesture in explicating visualisations as proposed by Roth and Pozzer-Ardenghi (2013) is reflected in the work of Kress and his colleagues (2001) and several studies in primary/elementary and junior high schools, the paucity of research in senior high school, especially in relation to complex sub-microscopic representations and comparison of different visualisations of the same phenomena, leaves open the question of the extent to which senior high school teachers are able to use such orchestration of these kinds of visualisations naturally and effectively to support students' learning. In this paper, we provide some initial insight into these issues in an examination of the use of gesture and language by two senior high school biology teachers in their explication and comparison of different types of complex visualisations depicting aspects of mitosis (cell duplication).

Data

The data is from two schools participating in the project investigating multimodal literacy development in senior high school science to address disadvantage in low SES areas (Unsworth et al., 2021). Both schools are assessed as close to the mean value of socio-educational disadvantage as measured by the Australian Curriculum Assessment and Reporting Authority (ACARA). The schools have 48% and 52%, respectively, of students with a language background other than English. The data was extracted from video recordings

of the teaching of mitosis by biology teachers of one year 11 class from each school. Both teachers, Ken and Tia, were in the early stages of their careers (< 5 years' experience) with similar specialist biology qualifications. We selected the video segments based on the difference between classrooms in the complexity of the diagrams of closely related phenomena and Tia's comparative explication of two complex diagrams.

Theoretical and Analytical Framework

Systemic Functional Semiotics

This study adopts the theoretical and analytical frameworks of Systemic Functional Linguistics (SFL) (Halliday, 1978; Halliday & Matthiessen, 2004; Martin, 1992) and its adaptations to image (Kress & van Leeuwen, 2006) and gesture (Ngo et al., 2021; Martin & Zappavigna, 2019). According to SFL, every instance of language simultaneously conveys three kinds of meanings: ideational (i.e. representing experience), interpersonal (i.e. enacting relationships) and textual (i.e. organising texts) (Martin & Rose, 2007). These three kinds of meanings are also communicated through the grammar of visual design in images (Kress & van Leeuwen, 2006) and gesture (Ngo et al., 2021; Hao & Hood, 2019; Martin & Zappavigna, 2019). We focus on ideational and textual meanings. Ideationally, we examine the linguistic, imagic and gestural representations of the scientific entities, their composition and activity. Textually, we show how the identification of entities and locations are followed or 'tracked' through the multimodal explications within and across different types of representations by the teachers' use of deictic gesture and verbal reference items using pronouns such as 'it', 'they', 'this', 'those' and/or demonstrative adverbs such as 'here' or 'there'. When these refer to an entity introduced within the language, it is called 'endophoric' reference. For example, 'We can see the spindle fibres coming out. They're not labelled ….'. Entities can also be tracked via *ellipsis* (Martin & Rose, 2007, p.179), which is marked ∅ with what is left out inserted in square brackets [] ('Each centriole is a ring of nine groups of microtubules….you can see here – there should be nine ∅ [groups of microtubules'].

These reference items may also refer to objects not previously mentioned. For example, a teacher may point to entities in an image and say, 'These inevitably shoot out', where 'these' refer to image entities not previously mentioned. This is called exophoric reference. Where exophoric references like 'these' or 'this' refer to entities whose identity is only retrievable from the material context, deictic gestures clarify what is being referred to.

Analysis

Our multimodal analysis is based on Baldry and Thibault (2006). The teacher's language was transcribed into segments aligned with co-occurring gestures shown as screen shots. The video segments were viewed multiple times to analyse the gestures and relationships among the language, gesture and image elements. Endophoric and exophoric verbal reference items are cohesive ties (Halliday & Hasan, 1976). The ties and their referents in the language and image respectively were tabulated along with descriptions of co-occurring deictic (textual) and ideational gestures. The analyses were initially conducted by author one and re-analysed by author two. Minor discrepancies (e.g. clarification of image referents for exophoric references) were resolved through discussion.
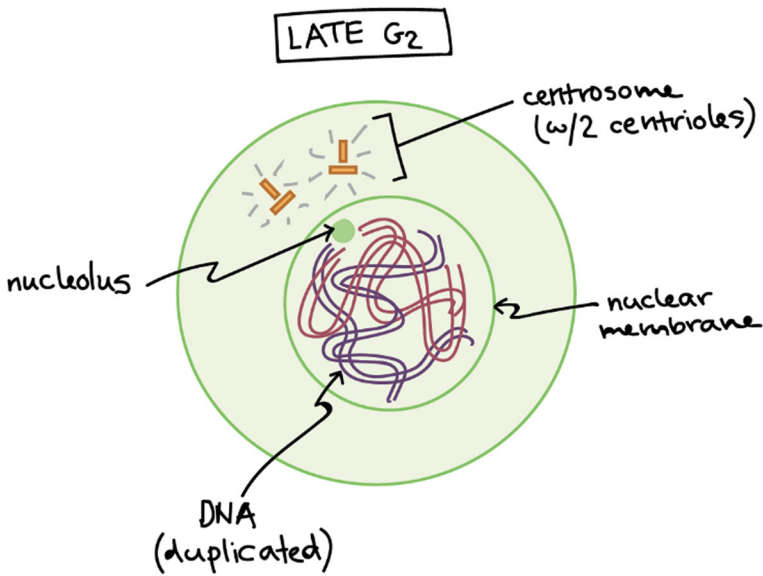
Fig. 1 Late interphase diagram

Findings

The videos from both teachers deal with the beginning of mitosis (late interphase/early prophase), emphasising the components of the cell, particularly the centrioles, and the activity of the microtubules/spindle fibres. Ken used the image in Fig. 1 concentrating on components such as the nuclear membrane, centrioles and spindle fibres.

Tia projected the images shown in Figs. 2 and 3 onto the screen simultaneously and focussed on the components of the centrioles and later the activity of the microtubules.

Whereas Ken identified the centrioles and their location, Tia's presentation involved greater depth and complexity of componental relations. Tia also sought to relate the two different representations of the centrioles, drawing attention to the differing affordances of the
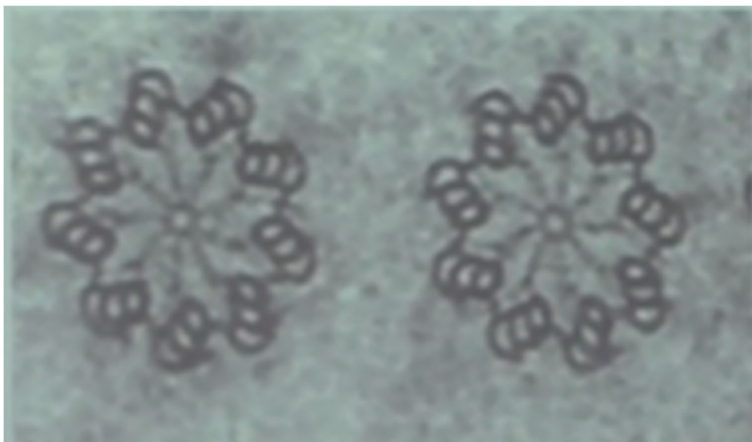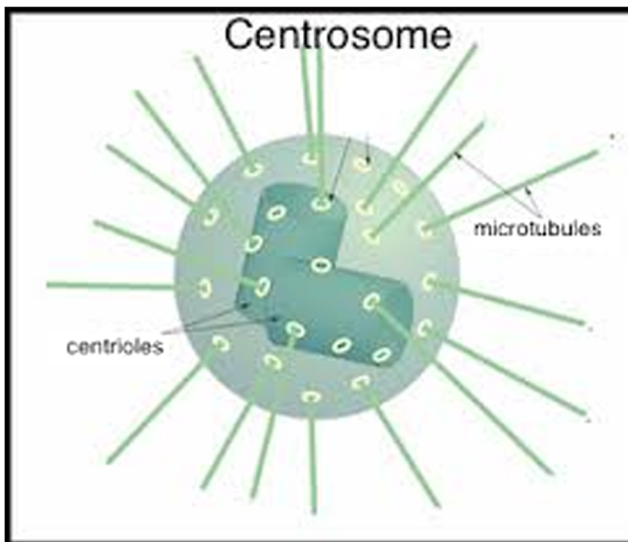

Fig. 2 Micrograph of a pair of centrioles

Fig. 3 Diagram of centrioles in the centrosome

micrograph in Fig. 2 and side view of the centriole in Fig. 3. We present, firstly, our analysis of two brief video excerpts where Ken identifies, locates and describes the nuclear membrane as depicted in Fig. 1 and then identifies the spindle fibres, their activity and location. Then, we present our analysis of three video excerpts where Tia (1) introduces the centriole components as depicted in the micrograph (Fig. 2); (2) compares the centriole views of the micrograph and the diagram (Fig. 3); and (3) focusses again on the micrograph, augmenting the image with her account of the microtubules' activity.
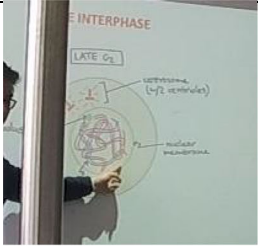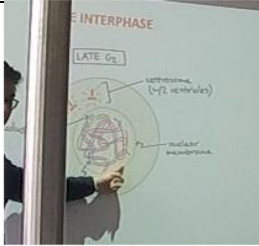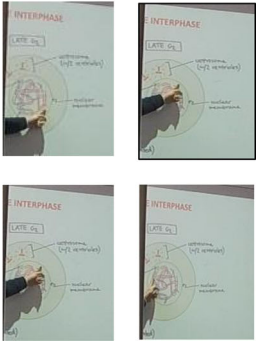
Explication of Late Interphase Mitosis Diagram

In EXCERPT K1, Ken identified the nuclear membrane as enclosing the DNA. Table 1 shows Ken's complete utterance and then each utterance segment and the accompanying gesture. The speech that co-occurs with gesture is bolded. In SEGMENTS 1 and 2, he emphasised the textual function of his pointing by sustaining the gesture throughout, and in SEGMENT 3, he used tracing around the circumference of the inner circle representing the nuclear membrane. This latter gesture not only textually identifies the relevant part of the image but also ideationally converges with the meaning of 'encompasses and surrounds'. In this excerpt, it is not necessary for Ken to use pronouns or demonstrative adverbs to refer back to entities in the language or the image as he clearly gesturally identifies the image component with the explicit verbal description of it.

The analysis of EXCERPT K1 is summarised in Table 2, where the SEGMENT shows language co-occurring with gesture in bold, the first mention of an entity in italics (*The nuclear membrane*) and ellipsis in parentheses. The identity of the referent may be its mention in the language and/or its image representation, and the cohesive tie is the language element that links to the referent within the language or image (Halliday & Hasan, 1976). This analysis demonstrates highly integrated orchestration of language and gesture in describing the nuclear membrane, its location and shape.

In EXCERPT K2, as shown in Table 3, Ken identifies the spindle fibres, their activity and location.
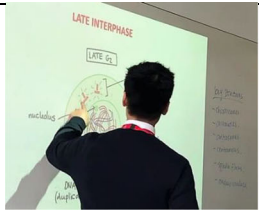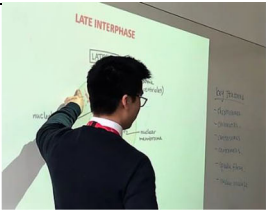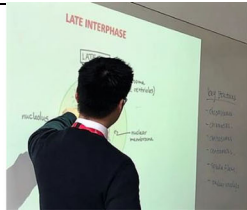
Table 1 Multimodal orchestration - EXCERPT K1

So, a couple of other things. The nuclear membrane is the case of the membrane that encompasses and surrounds the DNA material.		
SEGMENT 1	SEGMENT 2	SEGMENT 3
LANGUAGE: *The nuclear membrane*	LANGUAGE: *is the case of the membrane*	LANGUAGE: *that encompasses and surrounds* the DNA material.
GESTURE: sustained pointing at the section of image representing nuclear membrane throughout the language segment.	GESTURE: sustained pointing at the section of image representing nuclear membrane throughout the language segment.	GESTURE: tip of index finger tracing around the inner circle representing the nuclear membrane.

Table 2 Analysis of language and gesture - EXCERPT K1

SEGMENT	REFERENCE				GESTURE	
	ENDOPHORIC		EXOPHORIC			
	COHESIVE TIE	REFERENT LANGUAGE	REFERENT IMAGE	COHESIVE TIE		
So, a couple of other things.						
1 *The nuclear membrane*			Nuclear membrane		Points index finger	Sustained
2 **is the case of the membrane**	The membrane	Nuclear membrane	Nuclear membrane		Points index finger	Sustained
3 **that encompasses and surrounds** *the DNA material*			Nuclear membrane		Traces tip of index finger around perimeter	

Table 3 Multimodal orchestration - EXCERPT K2

So, we can see the spindle fibres coming out – they're not labelled, but they're there - coming out from there.		
SEGMENT 1	SEGMENT 2	SEGMENT 3
LANGUAGE: So, we can see **the spindle fibres coming out.**	LANGUAGE: **They're not labelled, but they're there.**	LANGUAGE: *coming out from there.*
GESTURE: repeatedly tapping on section of image representing spindle fibres	GESTURE: holding the pointing gesture on the same section of the image representing spindle fibres	GESTURE: holding the pointing gesture and repeatedly tapping on the section of the image representing the centrioles.

Table 4 Analysis of language and gesture - EXCERPT K2

SEGMENT	REFERENCE				GESTURE	
	ENDOPHORIC		EXOPHORIC			
	COHESIVE TIE	REFERENT LANGUAGE	REFERENT IMAGE	COHESIVE TIE		
1 **So, we can see *the spindle fibres* coming out**			Spindle fibres		Taps index finger	Repeated
2 **they're not labelled but they're there**	They	Spindle fibres	Spindle fibres	They	Points index finger	Maintained
	They	Spindle fibres	Spindle fibres	They		
			Spindle fibres	There		
3 **(They are) Coming out from there**	(They)	Spindle fibres			Points index finger	Maintained
			Centrioles	There	Taps index finger	Repeated

In all three segments, Ken used pointing gestures to identify the representation of the spindle fibres and the location from which they emanated. In each segment, he employed a different gestural strategy to emphasise the intended identification: increased the frequency of gesture (repeatedly tapping) in segment one, increased duration (holding the gesture throughout the speech) in segment two and the combination of these in segment three. The spindle fibres were introduced in his language in segment one while Ken was repeatedly tapping on their image representation. In segment two, he links 'the spindle fibres' in segment one to segment two using the pronoun 'they' (endophoric reference). But these pronouns are simultaneously referring out of the language (exophorically) to the image as Ken maintained his pointing gesture on the spindle fibres. In this way, Ken is sustaining the linking of the scientific terminology to the relevant part of the image. However, in segment three, although labelled in the diagram, he does not name the centrioles as the source of the spindle fibres but uses 'there' referring out of the language (exophoric) to indicate the location of their emanation, emphasising this by holding his pointing gesture and tapping repeatedly on the centriole.

This analysis is summarised in Table 4. Gesture links the image representation directly to the referent (as with 'spindle fibres' in SEGMENT 1). This linking also occurs simultaneously with an endophoric and exophoric reference (as with 'they' in SEGMENT 2). Apart from his not articulating the name of the centrioles, Ken's explication of the image in this excerpt shows a highly integrated orchestration of language, image and gesture, with his use of gesture establishing the convergence of the ideational meanings in his language with those in the image. Interestingly however, neither his language nor his gesture elaborates on the nature of the spindle fibres activity 'coming out'.

In both excerpts, there were more textual deictic gestures (pointing, tracing and touching using the index fingertip) than ideational gestures, while some functioned both textually and ideationally. The deictic gestures were emphasised through duration and frequency (tapping) as well as sustained tapping.
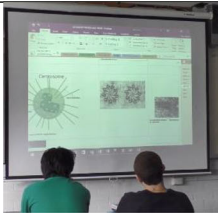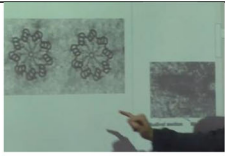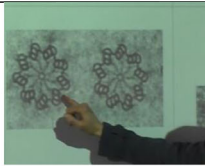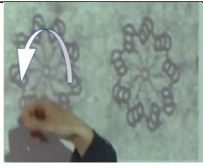
The analysis shows consistent co-occurrence of an identified entity or a specific location in the language (e.g. there) with a pointing gesture, and the co-occurrence of a tracing gesture with the description of the dimensionality of an entity (e.g. encompasses). Overall, there was a high degree of coordination of the verbal presentation and reference to entities, their dimensionality and location with gestures that clearly linked these to their image representations.

Comparative Explication of Two Representations of Centrioles

This section examines three excerpts from Tia's lesson focussing on the composition of centrioles and activity of microtubules. In EXCERPT T1, Tia sought to explicate the structure of the centrioles as represented in the micrograph (Fig. 2), which shows its nine component groups, each of three microtubules. Tia firstly read from the textbook and then addressed the class in relation to Fig. 2 projected onto a screen as indicated in Table 5.

After reading the textbook section, Tia's first two sentences signal her attention to transduction (Bezemer & Kress, 2008; Volkwyn et al., 2019)—relating the verbal textbook representation of the centriole structure to the micrograph. Her use of 'that' and 'It' refers back to the textbook section, and 'this' in sentence two refers to the projected images. Our analysis in Table 5 begins with Tia's third sentence as SEGMENT 1 in which she uses 'this' and 'here' to specify the micrograph, but she is still moving to the screen and there is no gesture yet to clearly identify the image to be focussed on. This becomes clearer in SEGMENT 2 as she prepares to point to part of the micrograph image, but the ellipsis, saying only 'nine'

Table 5 Multimodal orchestration - EXCERPT T1

Teacher reading from textbook:			
Each centriole is a ring of nine groups of microtubules. There are nine microtubules in each group [Student: three] three, sorry. There are three microtubules in each region group. Microtubules and centrioles are part of the cycle. In the complete animal cell centrosome, the two centrioles arrange themselves such that one is perpendicular to the other.			
Teacher's speech			
Now I want to look at what that looks like. It looks something like this. This is an electron micrograph. As you can see here – there are - should be nine. One, two, three, four, five, six, seven, eight, nine. This is a real image.			
SEGMENT 1	SEGMENT 2	SEGMENT 3	SEGMENT 4
LANGUAGE: This is an electron micrograph. As you can see here…	LANGUAGE: there should be nine (triplets of microtubules).	LANGUAGE: *One, two, three, four, five, six, seven, eight, nine.*	LANGUAGE: *This is a real image.*
GESTURE: None	GESTURE: None (only preparation move).	GESTURE: Little finger touching on each of the triplets as the teacher counted.	GESTURE: Little finger tracing around the image.

instead of 'nine triplets of microtubules', relies on the distant reference back to the textbook to connect the technical naming therein with the visually represented entities. In SEGMENT 3, Tia's precise gesturing with her little finger touching each of the nine textually locates them, and ideationally, coordinated with her counting, indicates the number of components without naming them. In SEGMENT 4, the exophoric reference 'this' co-occurs with the gesture of tracing around the perimeter of the micrograph image, thus identifying visually what is being classified as 'a real image'. The analysis of EXCERPT T1 is summarised in Table 6.
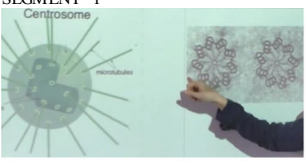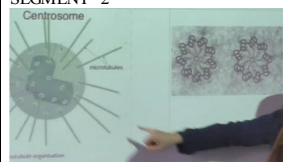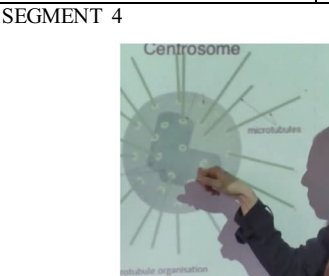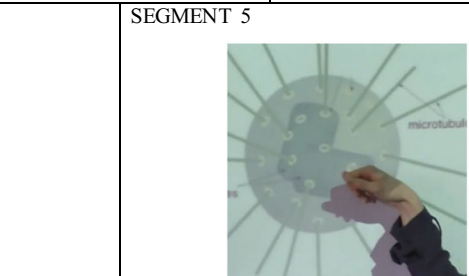
In EXCERPT T2, Tia compared the cross-section of the centriole in the micrograph (Fig. 2) with the side-view diagram (Fig. 3). Her orchestration of language, gesture and image is outlined in Table 7.

In SEGMENTS 1–3, there is no coordination of language, gesture and image in comparing the views of the centriole in Figs. 2 and 3. In SEGMENT 1, there is no verbal reference to any entity in Fig. 2, but Tia's gesture traced around the centriole. In SEGMENT 2, only 'this' was used to refer to the same entity in Fig. 2, without naming the entity (such as *this cross-section*

Table 6 Analysis of language and gesture - EXCERPT T1

SEGMENT	REFERENCE				GESTURE
	ENDOPHORIC		EXOPHORIC		
	COHESIVE TIE	REFERENT LANGUAGE	REFERENT IMAGE	COHESIVE TIE	
1 **This is *an electron micrograph*** **As you can see here**		Electron micrograph		This Here	
2 **there should be nine** (triplets of microtubules)					
3 **One, two, three, four, five, six, seven, eight, nine.**			Each triplet of microtubules on left centriole		Little finger touches each triplet
4 **This is *a real image*.**			Left centriole	This	Little finger circles perimeter

Table 7 Multimodal orchestration - EXCERPT T2

You can't really see this here. You can't really see that. What you're visualising is this.		
SEGMENT 1	SEGMENT 2	SEGMENT 3
LANGUAGE: ***You can't really see***	LANGUAGE: ***this***	LANGUAGE: ***here***.
GESTURE: little fingertip tracing over the centriole on the left of Figure 2.	GESTURE: Little finger is moving and pointing towards Figure 3.	GESTURE: little fingertip tracing over the circle representing the centrosome.

SEGMENT 4	SEGMENT 5
LANGUAGE: ***You can't really see that.***	LANGUAGE: ***What you're visualising is this***.
GESTURE: little finger pointing to the centriole in Figure 3.	GESTURE: little finger tracing back and forth at one area of the centriole image.

Table 8 Analysis of language and gesture - EXCERPT T2

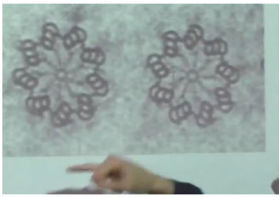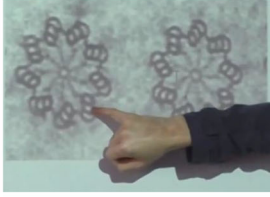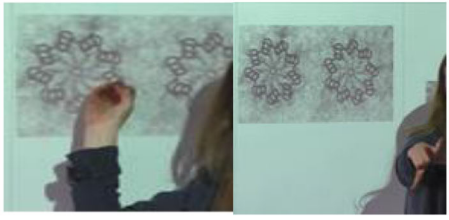| SEGMENT	REFERENCE				GESTURE	
	ENDOPHORIC		EXOPHORIC			
	COHESIVE TIE	REFERENT LANGUAGE	REFERENT IMAGE	COHESIVE TIE		
1 **You can't really see**			Fig. 2		Little finger circles perimeter	
2 **this**			Between images	This	Little finger points toward Fig. 3	
3 **here**			Fig. 3	Here	Little finger circles perimeter	
4 **You can't really see that**			Centriole in Fig. 3	That	Little finger points	
5 **What you are visualizing is this.**			Centriole in Fig. 3	This	Little finger traces section of centriole back and forth	Repeated back and forth

image of the centriole) and as Tia uttered 'this' she had ceased circling the centriole in Fig. 2 and was moving between the images with a pre-gesture action (pointing her little finger towards Fig. 3). In SEGMENT 3, Fig. 3 is introduced only by 'here' and a co-occurring deictic tracing gesture around the perimeter of the image, with no verbal description of the image. In SEGMENT 4, it is difficult to track back to what the pronoun 'that' refers. Co-occurring with this utterance is the gesture of the little finger pointing to the centriole in Fig. 3. Hence, this seems to be the referent. Logically, 'that' should refer exophorically back to the centriole image at the left of Fig. 2, but the deictic gesture does not support this, and the elliptical form of 'that' alone provides no disambiguation through naming of the entity (such as *the cross-section view of the centriole*). The ambiguity is exacerbated in SEGMENT 5, again through no explicit naming of the entity in Fig. 3 referred to simply as 'this' accompanied by the little finger tracing back and forth across a section of the centriole.

The analysis of EXCERPT T2 is summarised in Table 8. There is no ideational meaning realised through the language of science in the SEGMENT column. While there is a substantial amount of gesture, this is not coordinated with the exophoric reference to the images and since there is no verbal naming of the entities in the images, the comparison of what aspects of the centriole that can be accessed in the two images due to their different affordances remains unclear.

In EXCERPT T3 (Table 9), Tia used Fig. 2 in seeking to explicate the location of the microtubules and how they 'shoot out' from the centriole.

Following the preparatory SEGMENT 1, the gesture in SEGMENT 2 precisely traces the outline of each microtubule in the triplet with the tip of the little finger. As well as the deictic textual function of gesturally identifying the image entities, ideationally the gesture indicates their number and spatial dimensionality. In SEGMENT 3, 'these' and 'here' refer exophorically to entities and their location in Fig. 2 and co-occur with the gesture of tracing over the whole triplet of microtubules which identifies the image referents of the exophoric reference. However, there is no scientific naming of the entities referred to in 'these structures'. In SEGMENT 4, the absence of scientific naming despite the co-occurrence of the exophoric reference 'this' with the gestural tracing of the boundary of the entity and coordinating with the

Table 9 Multimodal orchestration - EXCERPT T3

As you can see, there are three-ah-three these structures here, and nine to make this an enclosed structure. These inevitably shoot out.		
SEGMENT 1	SEGMENT 2	SEGMENT 3
LANGUAGE: As you can see	LANGUAGE: *there are three-ah-*	LANGUAGE: *three these structures here*.
GESTURE: None (preparation for next gesture)	GESTURE: little finger traces the outline of each of the microtubules in the triplet	GESTURE: little finger circling over the whole triplet of microtubules multiple times.

SEGMENT 4	SEGMENT 5
LANGUAGE: **And nine to make this an enclosed structure.**	LANGUAGE: *These inevitably shoot out*.
GESTURE: Little finger circling the boundary of the centriole image on the left.	GESTURE: Fingers pinched together, moving from the visual text outwards to finish with a fully extended arm and fingers released to an open hand.

verbal description of 'an enclosed structure' leaves the multimodal orchestration incomplete. In SEGMENT 5, Tia's first three fingers pinched together and located on the microtubule triplet coordinates with the exophoric reference 'these', but again, the verbal referent is not provided. Next, the gesture of Tia's arm thrusting out from the image to full extension converges ideationally with the language 'shoot out'. The analysis of EXCERPT T3 is summarised in Table 10.

There are two main findings from the analysis of these three excerpts: (1) the scientific naming of the entities identified in Tia's introductory textbook reading was not reiterated in her explication of the visualisations; (2) deictic gesture and exophoric verbal reference to entities in the visualisations were frequently misaligned, especially in EXCERPT T2 comparing two visualisations.

Table 10 Analysis of language and gesture - EXCERPT T3

SEGMENT	REFERENCE				GESTURE	
	ENDOPHORIC		EXOPHORIC			
	COHESIVE TIE	REFERENT LANGUAGE	REFERENT IMAGE	COHESIVE TIE		
1 As you can see						
2 **there are three-ah-**			Triplet of microtubules		Little finger circles each microtubule in triplet	
3 **three** *these* **structures** *here*			Triplet of microtubules	These	Little finger circles whole triplet	Repeated
			Triplet of microtubules	Here		
4 **and nine to make this** *an enclosed structure*			Fig. 3	This	Little finger circles perimeter	
5 **These inevitably shoot out.**			Fig. 3	These	Fingers pinched together, move out from image to fully extended arm and fingers	

Discussion

Due to the paucity of research investigating senior high school teachers' orchestration of language and gesture in explicating sub-microscopic visualisations, there is a dearth of evidence as to how frequently or how effectively such orchestration occurs. The data from the early career teachers in this study may not be typical, but the analysis of their approaches points to issues for further research and for attention in the professional preparation as well as the ongoing professional development of teachers.

The single visualisation used by Ken (Fig. 1) is typical of such simplified diagrams of late interphase commonly occurring in senior high school biology textbooks as part of the sequence of diagrams representing successive phases in mitosis. His explication demonstrated highly coordinated, integrative deployment of language, gesture and image. While language is essential for naming scientific entities, meanings concerning location, shape and configuration are more appositely conveyed visually, due to the functional specialisation of images for conveying such meanings (Kress, 2003; Lemke, 1998). However, the naming and verbal dimensioning of the component entities need to be linked to these visual elements, necessitating deictic gesture co-occurring with exophoric references such as 'they' and 'there' or 'here'. Ken's deictic gestures of tracing or pointing were always clearly directed to the relevant image elements, sometimes emphasised through tapping and holding the gesture, and consistently co-occurring with verbal identification and dimensioning of those image elements—with the exception of his not verbalising the labelling of the centrioles as the source of the spindle fibres in EXCERPT K2: SEGMENT 3.

Ken's explication focussed on entities at the first level of a componential taxonomy of the cell at this phase of mitosis as represented in one simplified and commonly occurring image. A clearly more challenging orchestration of language and gesture was entailed in Tia's efforts to comparatively explicate two very different visualisations (including the less frequently encountered electron micrograph) representing, at a deeper level of the componential taxonomy, the components of centrioles. Our analysis showed the faltering coordination of language, gesture and image elements as she manoeuvred between the simultaneous display of the two visualisations in EXCERPT T2. The difficulty in co-deploying verbal reference and deictic gesture to track the same entities (albeit represented differently) across two very different types of visualisations was exacerbated not only because the scientific entities were not verbalised, but also because there was only one mention of image type (electron micrograph) and no discussion of the distinctive affordances of the different image types (Fredlund et al., 2014). Discussion of the representational affordances of the micrograph as a cross-sectional view and the side-view diagram may have helped to disambiguate some aspects of Tia's explication (EXCERPT T2: SEGMENTS 2 and 4) as well as advancing the students' metarepresentational competence (diSessa, 2004).

Both teachers made extensive use of deictic gestures, underlining their pivotal role in the multimodal orchestration of diagram explications. Ken's techniques for intensifying the identifying role of deictic gestures (tapping and sustaining the pointing position) may have assisted in clarifying the entities Tia was seeking to relate across visualisations. Both teachers also used gesture to augment the ideational meanings constructed by the images and the teachers' language. Ken traced the encompassing role of the nuclear membrane and Tia the activity of the microtubules. Her dramatic gesturing of how the microtubules 'shoot out' from the centriole augments the image representation of compositional relations with a bodily representation of activity that supports students' understanding of sub-micro processes by helping to re-imagine and re-size them mentally to relate to body-sized everyday processes (Gibbins et al., 2012).

The brief excerpts from these teachers' explications show them explicitly seeking to teach students how to interpret the 'constructedness' of visualisations, not only noticing what is represented but also what is omitted and how different forms of visualisation afford complementary representations of the same phenomena (Fredlund et al., 2014). By relating verbal, visual and bodily representations, they provide multiple perspectives on accessing different facets of meaning. In so doing the potential for enhancing student learning extends beyond the ephemerality of the teacher's portrayal at that moment to engaging and equipping students with an ongoing orientation to active probing and cross-referencing of visualisations. To realise this potential however, our analyses and the limited research in this area underline the need for more evidence about how teachers can most efficiently explicate complementary facets of meaning that can be accessed via the distinctive affordances of different kinds of complex, sub-microscopic and abstract representations of the same phenomena. Our analysis suggests a metalanguage describing different types of visualisations and their affordances as an important factor. Crucial to the efficacy of such metalanguage is ensuring the referents of demonstrative pronouns (e.g. this, these) and demonstrative adverbs (e.g. there, here) are made explicit verbally and their linkage to specific aspects of the visualisations clearly pin-pointed through effective use of gesture. Variation in the nature of complexity of the multiple representations that teachers use will put different kinds of pressure on the orchestration of language and gesture in their explication and it seems important to understand more about how teachers manage this.

Implications for Research and Teacher Professional Learning

Recent research has emphasised the significance of the interaction of gesture and language in strategic interpolations of teacher-focussed explication of visualisations as representation in teaching sequences shifts from enactive (action-based) to iconic (image-based) to abstract and symbolic representations (Tang, 2016, 2020, p. 136). The incorporation into such orchestrations of a metalanguage describing the meaning-making affordances of different kinds of visualisations is advocated by Tang (2020, p. 174); however, the studies drawn upon to support this are overwhelmingly located in primary and middle schools. Our study extends seminal studies on the role of gesture in science education by Roth and Pozzer-Ardenghi (Pozzer-Ardenghi & Roth, 2007; Roth, 2001; Roth & Pozzer-Ardenghi, 2013) by drawing attention to the importance of the synergistic deployment of such metalanguage with gesture in teachers' explication of complex sub-microscopic representations and their use of multiple such representations in senior high school science. Our data questions the assumption that the role of gesture in orchestrating intermodal meaning-making in diagrams is naturally utilised effectively by science teachers, whose professional preparation does not usually include a metalanguage for describing how language and other representational modes construct meaning. Key implications of our study are that student learning is likely to benefit if teachers are alerted to how their hitherto unconscious use of gesture can be effectively deployed to integrate the meanings that they convey through language with those they make accessible through other representations and if teachers are introduced to a metalanguage that assists them to convey to students, the semiotic affordances of different forms of representation and how these affordances influence our scientific understanding. Tang (2020. P. 174) has emphasised the significance of this for educationally disadvantaged students from low SES backgrounds. Opportunities for enhancing student learning would benefit from further research into senior high school teachers' use of metalanguage as part of their orchestration of language and gesture in their strategically deployed authoritative explication of multiple complex visualisations.

Declarations

Ethics Approval All participants have consented to participation in the research in accordance with the requirements of the Human Ethics Committees of the Australian Catholic University, Catholic Education Melbourne, and the Victorian Department of Education and Training.

References

Baldry, A., & Thibault, P. (2006). *Multimodal transcription and text analysis*. Equinox.

Bergey, B. W., Cromley, J. G., & Newcombe, N. S. (2015). Teaching high school biology students to coordinate text and diagrams: Relations with transfer, effort, and spatial skill. *International Journal of Science Education, 37*(15), 2476–2502.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication, 25*(2), 165–195.

Brandstetter, M., Sandmann, A., & Florian, C. (2017). Understanding pictorial information in biology: Students' cognitive activities and visual reading strategies. *International Journal of Science Education, 39*(9), 1218–1237.

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving students' diagram comprehension with classroom instruction. *The Journal of Experimental Education, 81*(4), 511–537.

Danielsson, K. (2016). Modes and meaning in the classroom–The role of different semiotic resources to convey meaning in science classrooms. *Linguistics and Education, 35*, 88–99.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293–331.

Fredlund, T., Linder, C., Airey, J., & Linder, A. (2014). Unpacking physics representations: towards an appreciation of disciplinary affordance. *Physical Review Special Topics - Physics Education Research, 10*(2), 020129.

Gibbins, I., Truman, C., & Ludy, M. (2012). A feeling for the image: Hands, body and visualisation of the invisible. In C. Kennedy & M. Rosengren (Eds.), *SPECTRA: Images and data in art/science* (pp. 129–134). Australian Network for Art and Technology.

Gilbert, J. K., & Treagust, D. (2009). *Multiple representations in chemical education* (Vol. 4). Cham :Springer.

Givry, D., & Roth, W. M. (2006). Toward a new conception of conceptions: Interplay of talk, gestures, and structures in the setting. *Journal of Research in Science Teaching, 43*(10), 1086–1109.

Halliday, M. A. K. (1978). *Language as a social semiotic: The social interpretation of language and meaning.* Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Halliday, M. A. K., & Matthiessen, C. (2004). *An introduction to functional grammar* (4ed ed.). Arnold.

Hao, J., & Hood, S. (2019). Valuing science: The role of language and body language in a health science lecture. *Journal of Pragmatics, 139*, 200–215.

Herrera, J. S., & Riggs, E. M. (2013). Relating gestures and speech: An analysis of students' conceptions about geological sedimentary processes. *International Journal of Science Education, 35*(12), 1979–2003.

Kottmeyer, A. M., Van Meter, P., & Cameron, C. (2020). Diagram comprehension ability of college students in an introductory biology course. *Advances in Physiology Education, 44*(2), 169–180.

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' ability to solve process-diagram problems in secondary biology education. *Journal of Biological Education, 49*(1), 91–103.

Kress, G. (2003). Genres and the multimodal production of 'scientificness'. In C. Jewitt & G. Kress (Eds.), *Multimodal literacy* (pp. 173–186). Peter Lang.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2ed ed.). Routledge.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: Rhetorics of the science classroom*. Continuum.

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). Routledge.

Márquez, C., Izquierdo, M., & Espinet, M. (2006). Multimodal science teachers' discourse in modeling the water cycle. *Science Education, 90*(2), 202–226.

Martin, J. R. (1992). *English text: System and structure*. Amsterdam: Benjamins.

Martin, J. R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause* (Vol. 1, 2ed ed.). Continuum.

Martin, J. R., & Zappavigna, M. (2019). Embodied meaning: A systemic functional perspective on paralanguage. *Functional Linguistics, 6*(1), 1–33.

Mathayas, N., Brown, D. E., Wallon, R. C., & Lindgren, R. (2019). Representational gesturing as an epistemic tool for the development of mechanistic explanatory models. *Science Education, 103*(4), 1047–1079.

Ngo, T., Hood, S., Martin, J., Painter, C., Smith, B. A., & Zappavigna, M. (2021). *Modelling paralanguage using systemic functional semiotics*. Bloomsbury Academic. (in press).

Nunes, T., Bryant, P., Strand, S., Hillier, J., Barros, R., & Miller-Friedmann, J. (2017). *Review of SES and science learning in formal educational settings*. University of Oxford/Education Endowment Foundation.

Padalkar, S., & Ramadas, J. (2011). Designed and spontaneous gestures in elementary astronomy education. *International Journal of Science Education, 33*(12), 1703–1739.

Pantidos, P., Herakleioti, E., & Chachlioutaki, M.-E. (2017). Reanalysing children's responses on shadow formation: A comparative approach to bodily expressions and verbal discourse. *International Journal of Science Education, 39*(18), 2508–2527.

Plummer, J. D., Bower, C. A., & Liben, L. S. (2016). The role of perspective taking in how children connect reference frames when explaining astronomical phenomena. *International Journal of Science Education, 38*(3), 345–365.

Pozzer-Ardenghi, L., & Roth, W. M. (2007). On performing concepts during science lectures. *Science Education, 91*(1), 96–114.

Roth, W.-M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research, 71*(3), 365–392.

Roth, W.-M., & Pozzer-Ardenghi, L. (2013). Pictures in biology education. *Multiple representations in biological education*, 39–53 Springer.

Ruiz-Gallardo, J. R., Fernández, B. G., & Jiménez, A. M. (2019). Visual literacy in preservice teachers: A case study in biology. *Research in Science Education, 49*(2), 413–435.

Siry, C., & Gorges, A. (2019). Young students' diverse resources for meaning making in science: Learning from multilingual contexts. *International Journal of Science Education*, 1–23.

Tang, K.-S. (2016). The interplay of representations and patterns of classroom discourse in science teaching sequences. *International Journal of Science Education, 38*(13), 2069–2095.

Tang, K.-S. (2020). *Discourse strategies for science teaching and learning: Research and practice*. Routledge.

Tang, K.-S., & Danielsson, K. (2018). *Global developments in literacy research for science education.* Cham: Springer.

Thomson, S., Bortoli, L. D., & Underwood, C. (2017). *PISA 2015: Reporting Australia's results*.

Treagust, D. F., & Tsui, C.-Y. (2013). *Multiple representations in biological education*: Springer Science & Business Media.

Treagust, D. F., Duit, R., & Fischer, H. E. (2017). *Multiple representations in physics education (10)*: Springer.

Ünsal, Z., Jakobson, B., Wickman, P. O., & Molander, B. O. (2018). Gesticulating science: Emergent bilingual students' use of gestures. *Journal of Research in Science Teaching, 55*(1), 121–144.

Unsworth, L., Tytler, R., Fenwick, L., Humphrey, S., Chandler, P., Herrington, M., & Lam, P. (2021). *Multimodal literacy in school science: Transdisciplinary perspectives on theory*. Research and Pedagogy: Routledge. (in press).

Volkwyn, T. S., Airey, J., Gregorcic, B., & Heijkenskjöld, F. (2019). Transduction and science learning: Multimodality in the physics laboratory. *Designs for Learning, 11*(1), 16–29.

Yeo, J., & Gilbert, J. K. (2017). The role of representations in students' explanations of four phenomena in physics: Dynamics, thermal physics, electromagnetic induction and superposition. In D. Treagust, R. Duit, & H. Fischer (Eds.), *Multiple representations in physics education* (pp. 255–287). Springer.