Even a two-year-old can do it!
The early stages of learning to understand moving-image media

Cary Bazalgette

Abstract. Film scholarship has consistently avoided discussing how we learn to understand the complex, multimodal systems of communication that moving-image media (referred to here as ‘movies’) have evolved into over the last 125 years. This article offers some reasons for this neglect: in particular, the popular assumption that movies are extremely easy to understand, and the relative lack of research on two-year-olds – the crucial phase in which this learning must take place. Drawing on a 20-month study of a pair of dizygotic twins, a vignette of their early viewing behaviour illustrates the features of focused attention which characterized their investment of energy in trying to make sense of movies. An analysis of this phenomenon, using concepts from embodied cognition, shows how instinctive responses relate to thought and reflection. Setting two-year-olds’ movie-watching within the wider contexts of story-reading, play and the enjoyment of repetition, the article provides evidence that such learning does take place and can be seen as a significant aspect of two-year-olds’ “entry into culture”.

Keywords: embodied cognition; moving-image media; two-year-olds; learning; emotion; viewing practices

It is strange that despite the huge volume of academic and critical writing about moving-image media that has flourished for many years, almost nobody has stopped to consider how and when we first learn to understand them. Much of this vast literature deals with the semiotic complexity and multimodal density of these media (for a selection of examples, see: Bazin 1967; Bordwell, Thompson 1980; Eco 1976; Frampton 2006; Keathley 2006; Metz 1974). References to the possibility of actually having to learn to understand them are rare. One scholar who has at least posed a question about this is Edward Branigan:

1 Culture, Communication and Media, UCL Institute of Education, UCL, London WC1H 0AL, UK; e-mail: carybaz@gmail.com.

https://doi.org/10.12697/SSS.2020.48.1.04
It seems remarkable that no one has undertaken to discover what special problems of narrative comprehension may be posed to a child by filmed narratives. For example, when and how do children understand an eyeline match, screen direction, cross-cutting, an unusual angle, off-screen space, or non-diegetic sound? (Branigan 1992: 225)

However, Branigan does not attempt to answer his question, which in any case only appears in a footnote. Other scholars have argued that the special problems he refers to do not really exist, or if they do, they are quickly figured out by children at such a young age that to investigate that process need not take up the time of serious scholars. In the preface to his book How to Read a Film, James Monaco asks “Is it necessary, really, to learn how to read a film?” and points out that “obviously, anyone of minimal intelligence over the age of four can – more or less – grasp the basic content of a film, record, radio or television program without any special training” (Monaco 1981: 17). Leaving aside his strange assertion that the four media he mentions present similar hermeneutic challenges just because they are all accessed on modern technologies, he justifies the title of his book by arguing that we do need to learn “how they tell us what they tell” – in other words, that learning how to read films properly is a task for older people: children can be satisfied with apprehending their basic content. On somewhat similar lines, Paul Messaris suggests that “research with children is of questionable relevance to the issue of a specifically cinematic literacy,” pointing out that “the children who participate in these studies are already veterans of many an evening spent in front of the tube” [i.e. the television] (Messaris 1983).

It has long been taken for granted that moving-image media (which from now on I shall refer to as ‘movies’) are easy to understand. From its earliest beginnings, the movie industry promoted itself as the most “lifelike” entertainment medium ever invented. The “founding myth” of movies is that viewers automatically believe they are watching “real life”: accounts of the first public screenings in 1895–1896 include the probably exaggerated claim that audiences ducked and shrieked en masse or even ran away when they saw a film of a train moving towards the camera (Loiperdinger 2004). Even now, many people still assume that movies (at least, live-action ones) are “a window on the world”. Since the 1920s, technological developments such as synchronised speech, colour, larger screens and immersive sound systems have become a normal feature of cinema exhibition and have all been driven by the desire to heighten the illusion that we are watching some kind of “captured reality”. But at the same time, developments in animation, computer-generated images, editing techniques and sound design have created pleasurably convincing fictional worlds that owe as much to fantasy, dreams and indeed nightmares as they do to real life.
Messaris presents a rather more persuasive argument about the “easiness” of understanding movies. He points out that many filmic devices, including for example eyeline matches, jump cuts and point of view shots, are in fact similar to people’s everyday perceptions and instinctive behaviour (Messaris 1994). These devices do not mimic such perceptions and behaviour exactly, but they do satisfy the desire to respond to visual phenomena such as a surprised or frightened gaze, a hand reaching out, or even just a moving vehicle. Experienced viewers know (or think they know) that the shot that comes after a shot of a gaze or a reaching hand will show what the characters are looking at or reaching for (i.e. a point of view shot), and that the shot of the moving vehicle will not necessarily be followed by other shots of the same journey, but by a cut to a shot of the vehicle arriving at its destination (i.e. a jump cut). The instinctive desire to find out the object of the gaze, the objective of the reaching hand or the destination of the vehicle is usually sharpened by our narrative expectations: our familiarity with movie narratives enables us to anticipate what is coming next. If we are wrong – for example if a surprised or frightened gaze is simply followed, for example, by the end credits of an episode in the TV series we are following, then our frustration may be tempered by eager anticipation, and speculation about what the character may have seen, while we wait for the next episode.

However, narrative experience is not the only factor at work here. Our desires and expectations in this context (as in similar real-life situations) are shaped by instinctive behaviour patterns that ensured the survival of our pre-human ancestors, and can be observed even in infants. Jerome Bruner describes how, even at nine months old, a child “looks out along the trajectory of an adult’s ‘point’ and, finding nothing there, turns back to check not only the adult’s direction of point but the line of visual regard as well” (Bruner 1990: 75). It is not difficult to figure out how essential this behaviour was for survival in hunter-gatherer societies, particularly when the point and gaze in question turned out to be indicating a hungry predator or some ripe fruit.

So why has there been so little interest in how and when children learn to make sense of movies? If we are to believe Monaco’s and Messaris’s claims that children are already expert viewers by the age of four, then they must be much younger when this learning process begins. In 2005, most UK babies were apparently watching TV by the age of three months (Marsh et al. 2005), from which we can assume that most of them were already listening to movie soundtracks while they were in the womb, and that these statistics would only differ today in terms of the increased diversity of platforms on which movies are now viewed. Yet some

---

movie-learning must take place between infancy and three or four years old, given that children in the latter age groups can usually follow most if not all of a complex, fast-moving mainstream feature film such as *Toy Story 4* (Cooley 2019).

Several factors contribute to the absence of serious attention to early opportunities for what I shall here term ‘movie-learning.’ Firstly, there is the popular perception of movies as easy to understand, as described above. Secondly, there is the recurring phenomenon of fears that new communication technologies and forms may present threats to the established social order, especially in relation to social categories that are seen as particularly vulnerable to seductively attractive sources of information or entertainment. Whereas this attitude used to focus on the dangers of reading for those deemed insufficiently armed with critical skills, such as the working classes or women (see for example Rose 2001; Wollstonecraft 1792), for more than a century it has been directed towards children and the technologies that deliver movies. In 1917 an independent inquiry on children and the cinema, commissioned by the National Council for Public Morals (a lobby group, not an official body, but with the backing of the cinema exhibition sector) concluded that “the cinema, under wise guidance, may be made a powerful influence for good; if neglected, if its abuse is unchecked, its potentialities for evil are manifold” ([Marchant] 1917: xxi). Since then, the research agenda on children and moving-image media has been dominated by a concern with the supposed risks (and to a far lesser extent – usually dependent on industry funding – the possible benefits) of watching, first of all films (until the 1950s), then television (until the 2000s) and now what, in the face of enormously proliferating formats, are rather helplessly summarized as ‘digital media’ or even just as ‘screens.’ There is research into what young children may learn from movies, but not into how they learn about them.

A major focus of the “risk” agenda has been the amount of time that children spend with moving-image media. This is reflected in everyday discourse, both face-to-face and online. It is widely assumed that television and other screen-based media are used as “child-minders”: that children are placed in front of them and kept quiet by the allegedly mesmeric effects of these media. Social media discussions of children’s movie-watching have many contributors who display a nervous sub-text about the ‘risks’ of too much ‘exposure’ – the latter being a term widely used by researchers (e.g. Christakis *et al.* 2004; Mar *et al.* 2010; Mistry *et al.* 2006; Stevens, Mulso 2006; Vandewater *et al.* 2005). The choice of the term ‘exposure’ is an interesting one, implying a complete denial of any process of interpretation being at work; it is an extreme version of the “movies are easy to understand” trope. The fundamental anxiety that underpins all these concerns is that time spent engaging with movies or “screens” will negatively affect children’s later ability to read, learn and concentrate; that it will model inappropriate
behaviour or induce mental health issues such as depression, and the more time spent on it, the worse the effects. There are now at least some critiques of this obsession with “screen time” but it has nevertheless penetrated everyday discourse to a remarkable extent (see for example Blum-Ross, Livingstone 2018).

A third factor that has deflected research and debate away from considering how children first learn to understand movies, is the remarkable lack of attention (commented on for example by Dunn and Woolley) that early years researchers have given to what I will refer to here as ‘two-year-olds’ (Dunn 1988; Woolley 1997). (I intend this to be understood here as a looser age-cohort than it looks, so I am including children in the later stages of their second year of life, and those embarking on their fourth, as well as those actually aged between 24 and 36 months; as a reminder of this I will also sometimes use the term ‘toddlers’ to avoid excessive repetition.) What research there is on two-year-olds tends to focus on the so-called ‘developmental milestones’ that mark the route to later learning, such as more fluent speech, socialization and increased dexterity, and factors that may negatively or positively affect these. Anxious parents are encouraged to look forward to these milestones, and to be patient with the challenges that their children are facing. There is thus an implicit ‘deficit model’ underlying the ways we tend to refer to children of this age, reflected in the way that we happily use terms such as ‘childish’ and ‘infantile’ to refer disparagingly to the behaviour of populist politicians, and to refer to this age-group as the ‘terrible twos’.

The ‘terrible twos’ concept seems to be confined to Anglophone cultures such as the UK, North America and Australasia, and is more adult-centred than child-centred with its implication that it is the carers who will inevitably have terrible problems with two-year-olds. Germans use the terms ‘Trotzphase’ or ‘Trotzalter’ (‘defiance phase’) which is still somewhat negative but at least more child-centred, and refers to a slightly wider age cohort (2–4): it derives from the word ‘Trotz’ which is an old word with a wide range of meanings from courageous resistance to stubbornness, spite and sulking. German developmental psychologists would use ‘Autonomiephase’, equivalent to the Danish term ‘uafhængighedsalderen’ (‘age of independence’), which relates more accurately to the period in which children start to achieve autonomy and self-affirmation. In Spanish, French, Italian, Turkish and Arabic (all I have checked out so far) there seems to be no term equivalent to ‘terrible twos’, although the wide circulation of English texts on child development and childrearing, in both online and print publications, has ensured that other cultures have started to pick up the English phrase as a way of describing the stage when children’s discovery of the words ‘me’ and ‘no’ presents parents with some interesting challenges.
Even a two-year-old can do it!

New parents therefore often approach the ‘terrible twos’ with trepidation (in Anglophone cultures at least), consulting the numerous social-media sites that provide advice on how to cope with tantrums, mood switches and insatiable curiosity. And it is when their children reach this age that parents also start to get anxious about ‘screen time’ child alone with a TV set and ‘digital technologies’ because of the effects that these are alleged to have on their children’s development. Many struggle to balance their fears of such effects against the convenience of having their toddlers sit still for half an hour with an iPad or TV set while household chores get done or dinner is cooked.

Nevertheless, two-year-olds are extraordinarily voracious learners, discovering an astonishing range of knowledge and skills. This is the period in which children learn to talk and to participate in their family’s social and cultural lives. Developing a sense of identity and personal agency, they never stop investigating and experimenting. But getting access to study them is not easy, because it usually involves having to study them in the home, and even when researchers just want to visit for, say, a few hours each month, that amount contact time is not enough to follow their development, to keep track of them as they move from one activity to another, or even just to understand their rapidly-evolving language skills. So it is not surprising that it is some of the scholars who have studied their own children who have been deservedly influential in the fields of education, child development, language and literacy (e.g. Britton 1970; Edmiston 2008; Halliday 1975; Piaget 1928; Weir 1970). However, two-year-olds’ movie-watching has rarely been studied, whether in their homes or elsewhere, and the studies that do exist (Briggs 2006; Robinson, Turnbull 2005) asked questions that were different from mine.

I undertook a 20-month observational study of toddlers’ viewing practices, looking for evidence that might enable us to describe their viewing as a learning process, rather than as the passive time-wasting deplored in popular discourse. I studied my own twin grandchildren (Connie and Alfie), and therefore did not face many problems of access or of understanding their speech. Through detailed analysis of more than 12 hours of video that showed their behaviour as they viewed a wide range of movies; through talk with their parents, the other grandparents and friends who also had two-year-olds, and by combing through social media, I assembled a repertoire of viewing behaviours that offer clues to toddlers’ interests and feelings about what they watch. I was, obviously, very much a participant observer, switching between my grandparenting and researcher roles as the situation demanded. In order to characterize the kinds of data I was gathering, I provide here a vignette of a viewing event which took place in the early stages of the research, when the twins were 23 months old, and which exemplifies some of their responses at that age.
Watching *Laughing Moon*

On a winter evening just two weeks before their second birthday, Connie and Alfie are roaming around their living room and browsing the toys, while their mother is in the kitchen getting dinner ready. Impulsively I reach for a DVD and choose a short movie that I know they have not seen before and that I think they might enjoy, even though it is very different from anything they have seen before: *Laughing Moon* (Nishimoto, 2000). As so often happened in this research, I had not been planning to do any observation that evening, but I decide to video their behaviour as they watch it. As usual, I use my iPhone for this, holding it as unobtrusively as possible, although the twins are used to being videoed and photographed and I know that they tend to take no notice of phones and cameras, unless they spot an opportunity to get hold of one and investigate it.

The movie begins, before they expect it to, with a loud ‘pop!’ and both children turn instantly to the screen. They see simple black geometric shapes against a white background, assembling into a figure something like a chicken, and hear a loud clucking on the sound track: they both freeze in mid-movement. Alfie is in the process of sitting down on a little wooden chair: he continues to grip the chair with one hand, while the other rests on the table in front of him as he lowers his bottom to the seat and watches tensely. Connie stands partly turned away from the screen: her attention is divided between it and a picture book that is lying face down on the table in front of her. She wants to try and open the book, which will not easily open backwards, and keeps flipping shut again whenever she lets go of it.

Both children maintain these positions until the movie ends, six minutes later. Alfie, highly sensitive to sound and easily frightened by sudden loud noises, maintains his awkwardly-seated position and his fixed gaze at the screen, jumping slightly when other loud sounds occur (e.g. rock music, dinosaur roar, gurgling water, motorcycle crash). Connie continues to try and open the book, but switches her attention to the screen whenever another sound attracts her attention. In these moments, her hands stay where they were (e.g. bending back the book cover, tweaking her sweater) but remain still, while she turns her head to the screen and watches intently. The movie offers a kind of game: the black “tangram” shapes keep reassembling as different characters while the soundtrack changes accordingly, and each character is teased by a little orange moon that keeps disguising itself (e.g. as a solid weight, a bouncy ball, a bird) and laughing as it escapes their grasp. Connie seems to get the point of this and begins to point, say “oh!” and glance at me when she recognizes a character (e.g. fish, dog). When the movie ends, both children immediately start to swarm around the room again, exclaiming “More! More!” I struggle unsuccessfully to operate the unfamiliar remote control, but then we are summoned to the kitchen for dinner and the viewing session is abandoned.
Focused attention

Anyone who has tried to persuade a toddler to stop investigating something really interesting like an electric socket or a lipstick in favour of a routine requirement like putting shoes on will be familiar with their capacity for totally focused attention. We all know what it feels like to be attentive to something, but if we actually observe and reflect on what a toddler is doing when she is really focused on some unfamiliar object or event, we can see how extreme their attentiveness can be.

For example, a toddler who wants to maintain an absolutely steady gaze at something that is not actually in his hands (a television screen for example) may well have to brace himself against a handy object such as a chair or even an adult's leg. Of course, as Jensenius et al. (2014: 208) point out, nobody can stand completely still, but given that toddlers’ centre of gravity is higher than that of older children and adults (Huelke 1998: 98), two-year-olds have to find ways of supporting themselves if they want to maintain steady visual contact, especially with a large area of moving images, as on a flat-screen TV set. So bracing is often just an essential response to keep the body stable. If there is no handy support, the child may stiffen his body and hunch his shoulders in the effort to maintain a steady position and may even have to pause now and then for a couple of seconds’ relaxation before resuming the rigid pose. An adult with an attentive toddler on their lap will be able to feel the child’s bodily tension and perhaps his grip on their limb or clothing as he maintains his gaze. Of course, if the child is also apprehensive about what he is looking at, the grip is likely to be tighter.

It is also interesting to observe what a toddler does with her hands, if they are not already in use as part of the “braced” posture. Wherever the hands were before her attention was focused, they are likely to remain in that position – so continuing to grip a bottle, cup or toy for example, or perhaps simply remaining placed on a nearby piece of furniture. There are parallels here with the way in which a predator such as a cat will “freeze” when it spots a movement that could be potential prey: if the cat is walking when this happens, one paw may remain raised so that no movement takes place that might alert the prey.

Besides posture, the other obvious physical feature of focused attention is facial expression. These can be extremely fleeting, and in the case of toddlers with their relatively plump faces, it can be difficult to spot some expressions such as a frown. Connie tended to chew her cheek when anxiously attentive, but it was only possible to notice this when it was captured on video, and had been recorded.

---

3 I am deliberately varying the gender pronouns I use in this paper, rather than opt for one gender.
when the lighting cast enough shadow to show the fleeting change in the shape of her cheek. More notably, both the twins tended to jut their jaws forward when they were not only concentrating but also a bit apprehensive. The major facial characteristic of more relaxed attentiveness in two-year-olds is commonly an open mouth, often accompanied by the typical toddler runny nose. Watching attentively for several minutes with one’s mouth open leads to dry lips, so lip-licking will happen regularly, and perhaps also a hasty nose-wipe with a handy sleeve. If a child is watching something attentively while drinking from a bottle or feeder cup, he may have to hold it to one side in order to maintain a gaze on the screen. Especially focused attention – in watching something suspenseful, for example – may be accompanied by deeper breathing.

Highly focused attention can often be observed when children see something new, so long as it appeals to them. Tastes will differ: the claim that “all children will love this” is, as we know, not necessarily reliable, and it is difficult – often impossible – to figure out what governs a two-year-old’s preference for one movie over another, just as her preferences for particular clothes, books, toys and foods may seem idiosyncratic. We can infer, however, that children’s preferences are based on important desires: in particular, recognition that the preferred object or experience has more to offer than can be discovered in a single encounter. Perhaps one of the most remarkable and fascinating things about two-year-olds’ learning is the fact that they make judgments about the learning-potential of an object or experience. We may disagree with such judgments if we think the object or experience is dangerous or inappropriate, but they can nevertheless be highly considered. The twins would usually watch a movie for several minutes before turning away or asking for it to be turned off, or they would continue to watch it through to the end and immediately demand to see it again. Laughing Moon was very different from anything they had seen before, and they were immediately intrigued by it. Two weeks later they watched it again, still a little bit tense about the loud noises, but now attuned to its game-like structure of presenting a series of different figures and using the soundtrack to add clues for identification, so that the viewing was punctuated by pointing and naming.

Movies, emotions and embodied cognition

The growing body of scholarship in neuroscience that traces the role of emotion in making sense of perceived phenomena shows how focused attention – and the indicators of tension that may go with it – can be read as the physical evidence of the complex but almost instantaneous processes at work in people’s brains.
as they make sense of what they perceive through their physical senses. Many film scholars now acknowledge the neuroscientific basis to the proposition that emotions play a key role in driving our interpretation of movies, as they do our thinking and actions in general. Recognizing and investigating the ways our physical perceptions and our evolved, instinctive responses relate to our cognitive processes is an important principle in embodied cognition: a theoretical field on which I drew substantially in my data analysis.

Several neuroscientists (e.g. Damasio 2006[1994]; Panksepp 2004) have discussed the role of emotion in cognition, extending the number of distinct emotions well beyond the five or six that are usually cited as ‘the emotions’, and explaining how they interact with motor actions and conscious thought. The ‘seeking’ concept as described by Jaak Panksepp seemed particularly relevant to the intense engagement I was observing in the twins as they watched the movies that interested them. Panksepp describes four “basic” emotions that must have existed in pre-humans: fear, anger, sorrow and seeking. While the first three of these obviously relate to the environmentally dangerous lives and close social interdependence of pre- and early humans, ‘seeking’ can be recognized as equally important for survival. Essential to feelings of engagement and excitement, it generates anticipation and investigation, not only in doing things like foraging and finding shelter, but also “gradually helps cement the perception of causal connections in the world and thereby creates ideas” (Panksepp 2004: 144–149). Panksepp does not refer to any physical signs that would indicate the presence of this emotion, but Damasio’s account of consciousness, citing wakefulness and attention, encouraged me to see ‘focused attention’ as a key sign of ‘seeking’:

Consciousness results in enhanced wakefulness and focused attention, both of which improve image processing for certain contents and can thus help optimize immediate and planned responses. The organism’s engagement with an object intensifies its ability to process that object sensorily and also increases the opportunity to be engaged by other objects – the organism gets ready for more encounters and for more-detailed interactions. (Damasio 2000: 182–183)

This offers a way of interpreting two-year-olds’ attentive movie-watching. A child’s intense attention to what is considered an appropriate task, such as working out how to solve a puzzle or do up their coat buttons, is admired; attentiveness to a movie is frequently labelled as ‘passive viewing’ or ‘addiction to TV’ – as a search for ‘toddlers and TV’ on parental social-media sites will quickly reveal.

---

4 Damasio (2000: 318–319) does not mean visual images, but “neural patterns or maps based on the momentary selection of neurons and circuits engaged by the interaction”.

---
However, Ofcom’s most recent data on children’s media consumption indicates that parental anxieties about their toddlers being addicted to TV either have little effect on their actual viewing practices, or that it is actually a small but vocal minority that expresses such anxieties. Ofcom’s annual research reports on children’s and parents’ media use and attitudes have only recent started including studies of three-year-olds; unfortunately, they still do not include two-year-olds. In 2018 they found that 96% of three-to-four-year-olds watched TV on a TV set for an average of 14 hours per week, while 30% also watched TV on other devices, mainly on a tablet. 32% watched TV programmes via what are called ‘over the top’ services, such as Netflix, Now TV or Amazon Prime Video. 52% of three-to-four-year-olds went online for nearly nine hours a week – much of which would have entailed going to YouTube for animated movies, funny videos or pranks. While most of the game-playing and social-media usage increased substantially for older age groups, it was still the case that watching movies on a TV set, although declining slowly but steadily overall, was still an important activity for three-year-olds.

We thus have to make inferences about the viewing practices of younger children. Although busy parents and carers increasingly give their toddlers a tablet or smartphone to distract them in public places such as a supermarket or a restaurant, at home they are more likely to leave a child alone with a TV set if she is unsupervised. So it may be reasonable to infer that the percentage of two-year-olds watching movies on a TV set is even higher than that of three-to-four-year-olds. There are thus many opportunities for two-year-olds to engage attentively with a movie, even though their family’s viewing practices overall may be more diverse. Where toddlers do have opportunities to watch together with others – adults and/or older siblings, or other toddlers at a childminder’s – this adds an important dimension to their movie-learning.

Ward (2015: 158) points out that “sound has the capacity to shape visual perception and steer visual attention”. As a sound designer himself, Ward takes his examples from cinema, but the children’s TV that the twins watched also included what he defines as “a process by which many sound fragments are created, selected, organized, and blended into a unified, coherent and immersive auditory image” (Ward 2015: 161). While I am not suggesting that the children could not tell the difference between sound from the screen and sound in the room, I do believe that, given toddlers’ acute awareness of what others do and say (Trevarthen 2005: 63–64), both verbal comments and emotive sounds from family members may contribute to their “reading” of a movie in the ways that Ward suggests when he comments on:

---
5 From recent interviews with parents of two-year-olds, and informal observation.
This is borne out by the way in which the twins themselves contributed to the “aural ensemble” of viewing events: pointing, calling out “oh!” and “ooh!” and naming characters, in a variety of tonal patterns. Many of these sounds – whether from adults or from children – became ritualized, as an expected part of the “social sound track” for movies they watched often. Wojciechowski and Gallese’s argument about “embodied simulation” may well be in play here. They argue that “by means of the neural format we share with other human beings, and, to an extent, with some animals, as well, we can map others’ actions onto our own motor system, as well as others’ emotions and sensations onto our own viscero-motor and somatosensory systems” (Wojciechowski, Gallese 2011). In this paper, they focus primarily upon neuroscience’s discovery of mirror neurons, established through finding links between visual images and cortical activity. What is important about this, they argue, is that the physical experience of simulating particular gestures, expressions or postures can give one an idea of the feelings that generate them. This phenomenon is exploited as a conscious technique in dance therapy and theatrical performance (Thom 2010), but in the quotation above and elsewhere, Wojciechowski and Gallese hint that embodied simulation may also be triggered by mechanisms other than visual perception. It certainly seemed to me that the twins’ responses to movies sometimes related to their awareness of the physical disposition and mood of the others in the room, both through bodily contact with them, and by hearing and registering at least the emotional tones of their comments and exclamations.

The embodied simulation concept must of course be playing a much more extensive role in two-year-olds’ movie-learning from the screen itself. All humans mimic other people’s expressions and gestures, often unconsciously, as part of normal social interaction. We tend to admire this as merely ‘cute’ when infants and toddlers do it, but it can more usefully be interpreted as a way of “trying out” the feelings that the bodily phenomena are expressing, and can relate to characters on screen as well as to real people. It may thus be a significant element of children’s learning to follow narrative, when they recognize, relate to and imitate the expressions, gestures and postures of characters in movies, and sometimes also those of people watching with them, and I observed many instances of this in my research. It is particularly relevant in the study of two-year-olds, given that their linguistic fluency and ability to follow narratives are still developing.
Repeat viewing

A typical aspect of toddlers’ movie-viewing behaviour is the desire to watch preferred movies – or sometimes just parts of movies – over and over again. Until the domestic video-recorder market began to expand in the 1980s, the only way anyone could watch a movie more than once was either to stay in the cinema until the show started again, or go back to the cinema another day and pay to watch it again – though some could also thread their home movies through their little 8mm film projectors and re-view them that way, and film scholars could of course watch as much as they liked on enormous Steenbeck editing machines. Now that the VCR is already nearly obsolete, and we can view and re-view on many different devices and platforms, children’s opportunities for re-viewing have greatly expanded, together with their range of digital operating skills. Children have always wanted to repeat experiences that are important to them: for example, playing a game over and over again or having the same picture book read to them many times, and re-viewing movies is no exception. We tend to comment on this by saying “she loves it” or “it’s his favourite”, but there is more to it than this: children need to repeat significant experiences until they have drawn out all of their meaning, and this can take a long time. It thus makes better sense to say, when a two-year-old does not want to watch a particular movie any more, that she has “used it up”, rather than “she’s bored with it”.

In one of her interviews with me, the twins’ mother, Phoebe, described their encounters with new movies: “I think when they’ve seen something for the first time they’re so awestruck by it that they tend to be very focused on it and then when they see it again they can start to predict what’s going to happen”. She implies that the children were using repeat viewings in order to establish narrative recall. It is certainly the case that with extremely familiar movies such as The Very Hungry Caterpillar and The Very Quiet Cricket (The Illuminated Film Company for Scholastic Productions, 1993), which they watched many times, they enormously enjoyed exercising their skills in recalling and predicting what was about to appear. Still, this is not the same as recognizing a causal chain of events, such as a character wanting to do or get something, or an event precipitating a further event or action, both of which rely on memory. In Bauer’s account of infant and toddler memory she points out that “the capacity for recall has been linked with the ability to provide a verbal report” (Bauer 2002: 137), but describes experiments involving imitation which show “that long-term recall processes are emergent by nine months, and that they become reliable over the second year” (Bauer 2002: 138). Although the children struggled to respond to the question “what happened?” – whether referring to movies or to real life – until around
Even a two-year-old can do it!

33 months of age, I believe that this was more a matter of lacking the vocabulary to express their interpretations fully than an inability to follow a causal chain of events.

Along with the surprising dearth of academic interest in the phenomenon of ‘repeat viewing’ that emerged in the 1980s with the introduction of domestic video players/recorders, little attention has been given to the movie-learning opportunities that repeat viewing may be offering. Marks Greenfield cautiously suggests that repeat viewing on video (she suggests “twice”!) could help with reading (Marks Greenfield 1984: 9), and Krcmar’s experiments indicated that repeat viewing helped 6–24-month-olds learn words (Krcmar 2010). There are clear parallels between the “routinized” opportunities for repeating “what others have said in previous recurrences of that situation” that occur during repeat viewings, and the function of such opportunities in language acquisition (Snow, Goldfield 1983). Repeat viewing parallels toddlers’ love of repetition in all sorts of contexts. For example, Bruner stresses the importance of the child’s social and communicative environment in the first two years of life, driven by what he terms “active means-end readiness” and “goal-directed activity”, present even in newborns (Bruner 1983: 24–27). Bruner also emphasizes the importance of parent-child games as a context for language learning (Bruner 1990). The children’s repeat viewing at this early stage (22–26 months) showed a combination of both factors: firstly, their eagerness to re-view, which they could express once they had ways of labelling the movies they liked; and secondly, opportunities to repeat words and phrases in a playful context with parents or carers.

However, the twins’ repeat-viewing practice continued well beyond this early stage. It is important to consider this in conjunction with the fact that, having repeat-viewed a movie for some time, they would often move on to something that offered new challenges. These would usually include a change of style and structure, but they could also involve factors such as more complex narratives, more dialogue and longer episodes. A significant example of this was their shift of attention at around 30 months of age – coinciding with considerable gains in verbal fluency – away from “edutainment” programmes such as The Numtums (Quinn, BBC 2012–2014) and Mister Maker (Morgan, Scott, Eyre, BBC 2007) towards Tree Fu Tom (Shaw, BBC 2012–2016), which they watched frequently for about four months. Making this change involved a leap from 5-to-10-minute viewings to 29 minutes, involving a mix of live action and animation, and plots that involved moral issues, such as dealing with the consequences of rash or self-interested actions, or defeating the machinations of malign characters. This was in a sense their “stepping stone” to becoming able to follow and enjoy full-length movies, although they would still sometimes like to re-view specific scenes from
movies, such as the “Step in Time” sequence from Mary Poppins (Stevenson 1964) and the “Sully putting Boo to bed” scene in Monsters, Inc. (Docter 2001).

Repeat viewing did not necessarily happen with everything the twins watched. In some cases, fearful or sad responses to a first viewing ensured that they avoided watching that movie again. Exploring the nature of these responses – often caused by what an adult would term a mistaken interpretation of what they were watching, especially in the case of what seemed to be unaccountable fears of quite innocuous scenes – formed a substantial part of my research, which I do not have space to recount here. I will simply note that a useful way of considering what happens in these events is to use Wojciechowski’s concept of diakresis. She adapts this from the work of Stanislas Dehaene on consciousness (Dehaene 2014). He argues that consciousness imposes a “bottleneck” on our capacity to process our perceptions: we have to select what is salient for us out of the mass of information constantly available to our senses (Wojciechowski 2015). Her term for this – ‘diakresis’ or ‘selection’ – can be adapted to help consider how two-year-old viewers deal with what they do not like in a movie. Wojciechowski’s adult viewers know (or think they know) what is meant to be narratively salient as they watch; two-year-olds may not: but this does not merely mean that they ‘make mistakes’ or ‘miss out’ on features that ‘really are’ salient to the narrative. What two-year-olds are undoubtedly doing as they watch with focused attention is to single out what is salient for them. Usually these are elements that they enjoy, but sometimes it can be something that characteristically distresses children of their age but disappears later: for example toddlers’ typical anguish about things that are broken or simply out of place (Kagan 1981). As they get older, their memory capacity increases; they see more movies and acquire wider knowledge of their culture, and they refine their ideas about salience.

**Understanding the nature of the system**

In Lesley Lancaster’s study of a two-year-old making a drawing with her father, she challenges the notion of ‘simplification,’ pointing out that it is “something socially and technically constructed” (Lancaster 2001: 134) – as we can see from the varied graphic styles in hundreds of other movies, books, toys, apps and packaging aimed at toddlers and pre-schoolers. Nevertheless, children seem to quickly assimilate and understand each stylistic trope on its own terms, with the help of adults (and, in the case of movies, also from voice-over commentary and/or dialogue) learning how to name each character, prop and setting. Lancaster’s quotation from Ferreiro and Teberosky’s account of children’s early reading is, I think, also pertinent to their early viewing:
Children pose deep questions to themselves. Their problems are not solved when they succeed in meaningfully identifying a letter or string of letters [here we could substitute “a shot or a sequence”], because they try to understand not only these elements or the results but also, and above all, the nature of the system. (Ferreiro, Teberosky 1982: 172)

Lancaster uses this insight to support her argument that two-year-olds’ interest in making sense of cultural forms is driven by their “expectations of significance”: they know very well that these forms are valued by others and they know they will be able to extract meanings from them if they try hard enough. And to do this, they need to understand the system of the cultural form they are engaging with. This also applies to the work that children must be doing when trying to make sense of movies.

Given the almost total neglect of pre-schoolers’ movies in the research literature – Buckingham (2002: 48), remarks that “what amuses and engages a two-year-old is, by definition, unlikely to hold much interest for us” – it is necessary to argue the case for the kinds of filmic complexity that two-year-olds may be encountering. Toddlers’ viewing material can be easily dismissed as “all cartoons”, and it is true that the websites of channels offering content for this age group are dominated by animated material and present relentlessly cheery, smiley and brightly coloured TV series. A children’s TV writer (interviewed by Steemers in 2007) pointed out that most “pre-school” television is actually pitched at “a generic four to five age range” (Steemers 2010). So even the moviemakers in this sector are making myriad choices from the sign systems at their disposal, which present children with fascinating interpretive challenges.

For example, children’s movies frequently play with filmic conventions. Many of these conventions serve to create and maintain the important concept of diegesis or “the story world”. But as Branigan (1002: 35) points out, this concept is a fuzzy one: if we accept that elements such as mood music, which we know are not experienced by the characters, are nevertheless “about the diegetic world [or ‘story world’] of the character and are meant to aid the spectator in organizing and interpreting that world and its events”; an alternative view might be that the diegetic/non-diegetic differentiation, though crucial, cannot always be sharply defined, and perhaps especially not in movies for children. Branigan nevertheless argues that “the spectator’s organization of information into diegetic and non-diegetic story worlds is a critical step in the comprehension of a narrative and in understanding the relationship of story events to our everyday world” (Branigan

---

This organizational process – whether it involves a sharp differentiation or a more nuanced reflection on possibilities – is likely to be an important one for children in making sense of the movies that they watch, given that these texts offer a very wide range of diegeses that are deliberately fanciful and governed by a wide variety of “laws”.

By mixing real-life and constructed settings, or live and animated characters, many children’s movies acknowledge and represent children’s capacity for establishing their own sets of rules for play scenarios. In his important book *The Work of the Imagination*, Paul Harris explains that pretend play starts in the second year of life, drawing on “the causal understanding of the physical and mental world that [toddlers] have already built up during infancy” (Harris 2000: 9). In pretend play, rules are established about what can and cannot happen, which are consistent within what Harris calls “the play framework” but do not necessarily correspond with the physical and social rules of the real world. Thus “the capacity to imagine alternative possibilities and to work out their implications emerges early in the course of children’s development and lasts a lifetime” (Harris 2000: xi).

A prosaic but significant, and, for a very young viewer, potentially puzzling feature of television’s relationship with the viewer is its extensive use of direct address – both visually, as when a presenter looks into the lens, and verbally, as when a commentary is either provided by a visible person or simply heard. Live programmes such as news presentations or sports commentaries, which address the audience directly and make efforts to draw us into the programmes through their introductory sequences, can also be seen as extending the story-world into our own living rooms, and encouraging us to think of ourselves as part of a much bigger audience, in many different settings. In addition, direct address, being apparently ‘live’, invites audience judgments that they are ‘real’ and perhaps ‘true’ even when, as often happens in children’s programmes, it is an animated character who addresses the camera – Peppa Pig for example, starts each episode with “Hallo! (snort) I’m Peppa Pig!”. Given that live presenters often have apparent eye contact with viewers and will use phrases like “see you tomorrow”, children may for a while believe that the presenters can actually see them, but soon develop alternative theories about this rather odd behaviour. For instance, Alfie at age 32 months speculated that the presenter of *Planet Jammbo* on the Milkshake channel, who spoke directly to camera and addressed the viewers as ‘Milkshakers’, was not speaking to him, but to “her children”, whom he could not see but who were probably, he said, “waiting outside”. By this age he was used to the demeanour of nursery teachers addressing a class of children and solved the ambiguity of the presenter’s role by identifying her simply as another of those adults who speak authoritatively to groups of children – who must therefore be somewhere nearby.
The channel hopes to convey viewers’ sense of solidarity with a wider viewing community by using the term ‘Milkshakers’, but the idea of ‘audience’ – as a larger group of people that you cannot see, and do not know – is quite sophisticated even for six-year-olds.

An acknowledgement of the specific needs of younger children emerged in 1997 when Ragdoll Productions were commissioned to create Teletubbies for the BBC, aimed at an audience aged between two and four. Although this stirred up much debate about its appropriateness, or lack of it, for this age group, and indeed about whether they should be watching television at all (e.g. Buckingham 2002; Linn, Poussaint 1999), this was the BBC’s most profitable programme in world markets, with 365 episodes backed up by huge merchandising (Briggs 2007). It was followed in 2007 by *In the Night Garden*, a 100-programme commission, marketed worldwide, also aimed at two-to-four-year-olds. As some academics (e.g. Bignell 2005) have noted, Ragdoll’s productions do offer quite complex interpretive challenges. Howard and Roberts’ observational study of children aged between 14 and 24 months watching *Teletubbies* (Howard, Roberts 2002) includes detailed analyses of the episode that they used, and confirms the cognitive challenges that required the children “to exercise developing theories of cause and effect, prediction and inference” (Howard, Roberts 2002: 334). The interesting feature of these enormously popular programmes is that they include deliberate appeals to the two-year-old audience, that challenge adults’ expectations of what a programme for little children should look like. For example, each *Teletubbies* episode includes a live action sequence showing some scene or activity calculated to interest two-year-olds, such as children playing or encounters with animals – which is always shown twice: adult viewers tended to find these sequences boring on first viewing, and excruciating when repeated, but repetition is essential for two-year-olds who are still learning to follow movie narratives.

A more complex feature is the deliberate disparities of scale which occur throughout *In the Night Garden* episodes. In some shots, the regular props and vehicles which appear in each episode appear as toy-size (i.e. only a few inches high); in others – often within the same sequence – they are shown as huge: big enough for the characters (most of whom are adult actors in bodysuits) to get inside them. The effect of this systematic breaking of scale-consistency is clearly intended to correspond to the efforts children make to enter the story-worlds of their play with toys: we may see a child lying down by a toy train track, for example, to attain the perspective of an inch-high protagonist, watching the train go by. As they create fictional play worlds, children often attempt to coordinate different scales: enabling dolls of different sizes to go inside a dolls’ house, for example.
The processes involved in figuring out the “rules” of each story world, and thinking about them, bear a close and important relationship to what is going on in much of two-year-olds’ play. Paul Harris points out that in pretend play, objective truth can be suspended in favour of make-believe truth. For example, in a game that involves pretending to wash a teddy-bear, the participants in the game will accept the “rule” that the teddy-bear really is wet (Harris 2000: 10). The development of play, he says, is an important way of developing understanding of fictional narratives: “it is not a distortion of the real world but an initial exploration of possible worlds” (Harris 2000: 27–28). So the proliferation of fantasy worlds in children’s movies offers children opportunities to extend these explorations, not, as is often assumed, confusing them about the difference between fantasy and reality. Harris argues that when we engage with fiction, we adopt a point of view that is inside an imagined world, accepting its “rules” and responding to its events as if they were real, but at the same time recognizing that they are make-believe. In 1817 the poet Samuel Taylor Coleridge coined the phrase “willing suspension of disbelief” to describe the mechanism through which we manage our engagements with fictional works of art (Coleridge 1817). So, both in their play and their engagements with fictional narratives – whether in books or movies – children become adept at suspending disbelief.

Conclusion

My research revealed an extensive range of viewing behaviours, only a few of which I have had space to discuss here. But even this limited selection demonstrates the intensity of two-year-olds’ commitment to a largely self-directed process of viewing and re-viewing as they figure out what movies mean. These findings challenge a number of dominant assumptions: that movies are easy to understand; that anything a toddler can learn is obviously extremely simple; that two-year-olds’ learning is not of great interest except insofar as it prepares for the later learning that we particularly value, such as speaking fluently, reading and writing, and behaving in socially acceptable ways. But in fact, learning to understand movies is now a vital part of how little children “come to learn about the culture they have been born into” (Trevarthen 1995: 7). It provides them with what are often self-directed contexts in which to investigate and eventually understand the key features of narrative, as well as the specific narration techniques that are unique to movies. Studies on this age cohort are relatively rare, and studies of their “movie-learning” are even rarer, but they can lead us to respect two-year-olds’ achievements in self-driven cultural learning, and perhaps to acknowledge them when we devise their subsequent school curricula.
Even a two-year-old can do it!

References


Bordwell, David; Thompson, Kristin 1980. *Film Art: An Introduction*. Reading: Addison-Wesley.


Even a two-year-old can do it!


Даже двухлетний ребенок справится! Ранние этапы обучения пониманию медиа «движущихся картинок» (moving-image media)

В киноведении последовательно избегали обсуждения того, как мы учимся понимать сложные, мультимодальные системы коммуникации, такие, как движущиеся картинки (называемые здесь «фильмами»), которые развивались на протяжении последних 125 лет. В этой статье приводятся некоторые причины такого пренебрежения: в частности, распространенное предположение, что фильмы чрезвычайно легко понять, и относительное отсутствие исследований двухлетних, т.е. референтной группы в возрасте, представляющим собой решающий этап, на котором должно происходить это обучение. Опираясь на 20-месячное исследование пары близнецов, мы обнаружили, что сценарий их поведения на ранних стадиях просмотра иллюстрирует особенности сосредоточенного внимания, которое характеризовало их энерговложение в попытку разобраться в фильмах. Анализ этого явления с использованием понятий из области имплементационного познания показывает, как инстинктивные реакции соотносятся с мыслью и рефлексией. Рассматривая просмотр фильмов двухлетними детьми в более широком контексте чтения историй, игры и наслаждения от повторения, в статье приводятся доказательства того, что такое обучение действительно имеет место и может рассматриваться в качестве важного аспекта «вхождения двухлетних детей в культуру».

Isegi kaheaastane saab sellega hakkama!
Varajased etapid liikuvpildilise meedia mõistmaõppimisel

Filmiuuringutes on järjepidevalt välditud selle käsitlemist, kuidas me öpime mõistma keerukaid multimodaalseid kommunikatsioonisüsteeme, mis on viimase 125 aasta jooksul välja kujunenud liikuvpildilises meedias (mida siinkohal nimetame ‘filmideks’). Artiklis pakutakse välja sellise unarussejätmise mõned põhjused: eelkõige laialt levinud oletus, et filmdest aru saada on äärmiselt lihtne, ning uuringute suhteline vähesus, mis puudutaksid kaheastaseid – otsustav faasi, mil selline öppimine peab aset leidma. Visandis, mis tugineb kahemunakaksikutate paari 20 kuud kestnud vaatlusele, illustreeritakse fokuseeritud joomi, mis iseloomustasid seda, kuidas nad pühendasid energiat püüetele filmidest aru saada. Selle nähtuse analüüsime kehandunud taju mõisteid kasutades näitab, kuidas instinktiivsed reaktsioonid suhestuvad mõtlemise ja refleksiooniga. Paigutades kaheaastaste filmivaatamise lugude lugemise, mängu ning korduste armastamise avaramasse konteksti, pakutakse artiklis tõendeid, et selline öppimine toimub ja et seda võib pidada oluliseks aspektiks kaheaastaste „kultuuri sisenemisel“. 