

Production Features as Scaffolds for Children's Learning:
Lessons for Instructional Design

Sandra L. Calvert

When Americans think of literacy, they generally think of children's ability to read and write words. This kind of literacy is indeed an important way to become a well-versed citizen in the information age (Neuman, Copple, & Bredekamp, 2000). However, often neglected and even diminished in importance are visual, iconic modes of thinking which tend to be prominent in the information technologies that comprise children's daily lives (Calvert, 1999). Visual and non-verbal icons are also a legitimate mode of thinking, and one that is undervalued and underused to reach children who do not readily understand the abstract verbal symbols that are required for success in school settings.

Media use a rich display of visual and auditory production features, known as formal features, that can readily be used to instruct and teach children even as they are being entertained (Calvert, 1999). In particular, formal features such as action, sound effects, and singings can be used as scaffolds, building bridges between how a child thinks at particular points in development to the knowledge to be learned.

Little is known about how children come to understand these visual symbols, about how children come to understand and become literate in the daily media experiences that pervade their daily lives. Nor do we understand enough about how these visual and non-linguistic auditory symbols provide links to verbal words and enhance or diminish literacy. Yet from the cradle to the grave, African American children, who are often economically disadvantaged, are more likely to live in homes that are television dominated than are their Caucasian peers (Roberts, Foehr & Rideout, 2005). While

television viewing is often blamed for poor literacy, there is the potential for television to enhance literacy as well.

In this chapter, I will pursue two major lines of research and thought. The first involves how formal features can be used to facilitate children's learning and understanding of verbal material, a more traditional kind of literacy. In this area, I will focus on the role that formal features can play as scaffolds for children's verbal learning. The second line of thought and research involves how children come to understand the unique codes of media as a symbol system in its own right. The implications for effective instructional design will then be discussed.

Formal Features

Formal features are audio-visual production features that structure, mark, and represent content (Huston & Wright, 1983). Formal features are the grammar of audio visual media (Calvert, 1999). At a macro level, these features involve *action* (physical movement) and *pace* (the rate of scene and character change). At a micro level, these features include visual camera techniques like *pans* where objects are followed in a continuous sweeping motion, *zooms* to close-ups or away from objects, *fades* where the screen goes black, *dissolves* where one object appears on top of another, *cuts* from one point of view to another, and *visual special effects* where the physical rules of reality are violated. Micro auditory features include *sound effects* in which loud noises occur, *character vocalizations* of unusual non-speech sounds, *music* that is either in the *foreground* (prominent) or *background* (with dialogue), *singing* (music and language combined), *laugh track* (off-screen audience laughs), *dialogue by adult* (adults speakers) *child* (child speakers), or *non-humans* (e.g., animals characters in cartoons), and

narration (one person speaks and explains the story). See Table 1. Being able to decode and “read” these features is the essence of being a literate user of information technologies.

Formal features vary in perceptual salience, that is, in their attention-getting properties. Perceptual salience involves stimulus properties such as movement, contrast, change, incongruity, and complexity which are likely to have survival value to our species (Berlyne, 1960). Applied to information technologies, macro features like action (movement), rapid pacing (change and complexity), and micro visual and auditory features like character vocalizations, sound effects, visual special effects, and frequent camera cuts (incongruity, change, contrast) have been classified as perceptually salient. By contrast, features like adult dialogue, child dialogue, and narration, while potentially carrying the most informative linguistic content, are in and of themselves low in salience.

The linguistic features of adult dialogue, child dialogue and narration have considerable potential to foster children’s interest in and sensitivity to oral language. If words are written on the screen, literacy skills can also be promoted. However, the varying salience of different features used to present content in information technologies creates a challenge: how do you get children to pay attention to and process language when there are so many other more interesting features to attract and hold their attention? In many media interfaces, the answer lies in the judicious pairing of other features- notably action, sound effects, character vocalizations, and singing- with written and spoken language. These features are maximally effective as a learning tool when they dovetail with the cognitive skills that children bring to bear on information at different points in development.

Formal Features as Scaffolds for Children's Learning

Vygotsky (1978) argued that young children needed scaffolds, or bridges, that link what children currently know and understand to knowledge that is just beyond their realm of understanding. His approach focused on verbal, linguistic ways of learning, particularly the bridges that parents build between what their child knows and the knowledge to be learned. Just as knowledge may be more or less available to children during certain developmental time frames, so too can children's abilities to process certain symbol systems. According to Bruner and his colleagues (Bruner, Olver & Greenfield, 1968), very young children think in enactive, motoric modes followed by visual iconic and finally abstract symbolic verbal modes of thought. These modes of thinking can be activated and enhanced by using certain production features in a presentation, providing scaffolds to improve the effectiveness of informal instructional learning environments, including those presented in a linguistic format. While less understood than the verbal scaffolds presented by parents, one thesis developed here is that formal features can also serve as scaffolds for children's learning.

Features as early elicitors of imitation and enactive ways of thought. Thinking with the body can readily be observed in infants' and young children's imitation of the many live adults and symbolic models that they view. As early as 14 months of age, Meltzoff (1988) demonstrated that infants can imitate actions portrayed on a video, though the presentation had to involve contingent replies to what the baby did. These early imitative activities activate the mirror neuron system in which babies come to know by copying their movements (Meltzoff, 2002).

We examined songs as ways to promote early learning through a scaffold of enactive rehearsal (Calvert & Goodman, 1999). Toddlers either sang songs or sang the songs while displaying actions that conveyed the meaning of the song. For instance, they either sang *I'm a Little Teapot* and tipped their bodies over and “poured out” the tea, or they just sang the song. The toddlers who used enactive ways of rehearsing the song lyrics subsequently understood the meaning of the song better than those who simply sang it without the aid of motor rehearsal. Put another way, the motor behaviors had provide a scaffold, i.e., a link, to the meaning of the song lyrics.

Similar findings occurred when preschool-aged children were exposed to an episode of *Dora the Explorer* in which we manipulated the interactive prompts that get children to act out content (Calvert, Strong, Jacobs & Conger, in press). In the original version, Dora asks children to do things with her like climb the ladder to rescue Benny the Bull, thereby eliciting imitative actions from her verbal prompts. We manipulated this prosocial, educational story so that children were exposed to either: 1) a control version where there was no interaction and an adult sat at the back of the room; 2) a control observational version where the adult sat beside the child while viewing; 3) a participatory version where the adult interacted with Dora and the child could do so or not; and 4) an interactive version where the program paused and the child had to use a computer to move the program to the next point. Children from Latino backgrounds, who tended to come from Head Start programs, were compared to their Caucasian peers who came from middle-class backgrounds. After viewing, children answered a verbal multiple-choice test. We found that the more children participated with Dora, the better they understood the central story content. Although the middle-class Caucasian children

understood the content better than the lower-income Latino children did, Latina girls who were in the interactive condition understood the central story content better than the Latina girls in the observational condition. The results suggest that participating with content- in which children enact behaviors that are relevant to the story- yields better comprehension than simply observing it.

Action as a mode of thought. Preschool-aged children often think in a visual iconic, mode (Bruner et al., 1968). Action, or movement, can enhance or disrupt children's learning of content, be it presented on a television or a computer (Calvert, 1999).

The classic study of children's learning from television was conducted by Hayes and Birnbaum (1980) in which they showed preschool-aged children original intact programs or altered programs that had the sound track of one program combined with the visual track of another program. In mismatched conditions, children remembered the visual track over the verbal track. This outcome became known as the visual superiority effect: children remember the visual track at the expense of the auditory track.

In our early work (Calvert, Huston, Watkins & Wright, 1982), we wondered if the visual superiority effect was really because of the perceptually salient character actions. Moreover, because action and dialogue are similar to iconic and symbolic modes of thought, we believed that action could facilitate children's learning, particularly at young ages when the predominant mode of thinking is more iconic than at subsequent points in development. By contrast, we expected that dialogue alone, a non-salient form, would be of more benefit to older children. In our initial correlational study, preschool-kindergarten and third-fourth graders viewed an episode of *Fat Albert and the Cosby*

Kids, a prosocial cartoon featuring a cast of animated African American boys and their live African American host, Bill Cosby. We had previously scored the formal features of that program, including the character action, and we later linked children's attentional patterns to the program and their comprehension of program content to those features. Children remembered the central plot-relevant content better when the language had been paired with action rather than when the central content was only spoken. These findings were true of the older as well as the younger children. In fact, the youngest children who understood the program best did not attend when Bill Cosby was speaking, presumably because his narrative comments occurred during low-action sequences and were incomprehensible so they stopped looking. Our results suggested beneficial effects of action through the middle childhood years. Similarly, 4 and 7 year olds reproduced the actions more than the language of characters when they heard or viewed a story (Gibbons, Anderson, Smith, Field & Fischer, 1986).

Lawler (1982), who was creating computer worlds like "beach world" during this same time frame, was able to improve his two-year-old daughter's literacy skills by building vocabulary words into the program. For the object to appear, his daughter Peggy had to key in the word. Then the object appeared, i.e., she had control of making events happen. For instance, if she keyed in the word "pony", a pony appeared on her beach. Lawler created what became known as intrinsically interesting learning environments.

In subsequent experimental studies, my colleagues and I decided to examine the role of various formal production features in creating intrinsically interesting learning environments. In our first study (Calvert, Watson, Brinkley & Bordeaux, 1989), we created a "park world" in which objects were programmed to appear with or without

action and with or without sound effects. Objects belonged to one of six groups (e.g., vehicles: car, truck, plane, and train), resulting in 24 total objects. Four versions of park world were created in order to examine the effects of movement and sound effects independently of the specific object. For example, across the four versions of park world, the car appeared moving with sound effects, moving without sound effects, nonmoving with sound effects, and nonmoving with no sound effects. Kindergarten-aged children listened to the same story about park world for four days as the experimenter keyed in the objects. Each day, children selected objects to go in their park from these 6 sets and keyed in those object names. Children chose more objects that moved than those that were stationary to go into their park. On the fifth day, we asked children to tell us all of the objects that were in park world. Objects that moved were more likely to be recalled than were those that were stationary. The findings suggested the importance of action for remembering information.

In our next study (Calvert, Watson, Brinkley & Penny, 1990), we randomly selected a version of park world and added a voice synthesizer to speak the object names, creating what became known as talk world. The action was retained. Here we became particularly interested in how to link action and language to improve children's recall of content. Children heard the story only one time in this version and were then asked to tell us the names of all the objects that they could remember from talk world. This time we compared good and poor readers at grades kindergarten and second. There were no effects for kindergartners who were not yet reading. However, poor readers in the second grade recalled just as many objects as the good readers when the objects moved. By contrast, good readers recalled more objects than poor readers when the objects were

stationary. The results suggested the beneficial effects of action for poor readers, as action basically provided a scaffold for children's verbal recall of object names.

Rehearsal is a mechanism that enhances children's learning of educational content. Following Salomon's (1974) ideas of activating (calling upon) versus supplanting (providing) cognitive skills, I (Calvert, 1991) modified how the story was presented in talk world. In the verbal label condition, we continued to read the words (i.e., supplant). In the no label condition, rather than reading the targeted words, we now paused (i.e., activation condition). Objects still moved or were stationary. Preschoolers' and kindergartners' spontaneous production and subsequent recall of objects was then examined. Children in the no label condition were far more likely to produce the names of targeted words, particularly when the objects also moved. For both age groups, using action and verbal labels together facilitated recall better than having no action or no label. The kindergartners' recall was most likely to benefit from having both verbal labels and action presented simultaneously, suggesting that they were able to integrate the different forms of presentation better than the preschoolers did. Here, then, we see beneficial effects of action as a scaffold for language for both rehearsal and memory of content.

Children with developmental disabilities like autism often are challenged when learning language, in part because they are more interested in machines than in the human interactions that are a major context for language development (Baron-Cohen, Wheelwright, Lawson, Griffin & Hill, 2002). Computers, therefore, are a potentially useful mechanical domain for teaching language to children with autism. To examine the potential of production features in a computer scenario for children with autism, Monique Moore and I created a scenario where children could learn words that were associated

with moving objects on a computer display. The Lovaas method, in which children are taught to pay attention to an adult to facilitate language acquisition, was used in conjunction with the computer program or alone as the control condition. After instruction, children were asked if they wanted to play outside or keep playing on the computer. We also assessed their verbal learning. Preschool children who had autism were more motivated to continue playing the computer game and learned more nouns when they had been instructed via the computer than by the Lovaas method alone (Moore & Calvert, 2000).

In summary, while visual content can interfere with children's recall of verbal content when a mismatch occurs between the verbal and visual content, action generally helps children understand verbal content when these two forms present consistent information. Beneficial effects occur whether the content is presented via television or computer interfaces. Action seems particularly helpful for children who are young, who don't read well, and who have developmental problems like autism. Even though older children get better at processing words without some kind of visual scaffold, action can still benefit children's learning well into middle childhood.

Singing for verbatim recall. Singing is a reflective feature that combines language with music (Huston et al., 1981). While educators have often thought that singing is a useful way to improve learning, the scaffold that it provides to language provides a bridge to verbatim memory, not to comprehension of content.

School House Rock was an instructional series designed to teach children history, science, mathematics, and English. Short vignettes of approximately 3 minutes presented animated bits with singing to accentuate the message. In our first study of School House

Rock (Calvert & Tart, 1993), we examined a vignette about the Preamble to the Constitution. College students were compared in conditions in which the Preamble was either spoken or sung. Immediate and long-term recall of the words to the Preamble was assessed after several viewings. Students who had seen the sung version of the Preamble recalled more of the words in the exact original order than those who had seen the spoken version of the vignette. The results linked repeated exposure to a sung vignette to very good short-term and long-term memory of words.

In a subsequent study (Calvert, 2001), I examined a history vignette titled *The Shot Heard Round the World* in conditions that varied the use of visual or a non-visual track and the use of a sung or a spoken audio track. This time, children viewed the vignette only once. Contrary to prediction, children understood the spoken track better than the sung track. While unexpected, these findings dovetailed nicely with an intriguing finding from our Preamble study: words were sometimes substituted in students' renditions of the Preamble that did not preserve the original meaning of the Preamble. For example, one person in the singing condition wrote "...to ensure the blessings of liberty to ourselves and our *prosperity*..." rather than "...to ensure the blessings of liberty to ourselves and our *posterity*..." (italics added). These findings suggested some problems in using singing to teach information that went deeper than superficial memorization of the lyrics.

I'm Just a Bill, in which repetition was manipulated, was the focus of another aspect of this study (Calvert, 2001). The bill was depicted as a piece of paper with writing on him; he went through the process of becoming a law during the course of the vignette, which we presented in its original spoken and sung soundtrack. Third graders and

college students were exposed to the vignette either once or four times. After the final exposure, verbatim recall, verbal sequencing of how a bill becomes a law, and comprehension of the content was assessed. The repetition condition increased all students' verbatim recall and verbal sequencing of content, but not their recognition of the important story content. When asked what a bill was, one child told us that a bill was "something that you pay."

Taken together, the findings suggest that singing provides an excellent and durable way to rehearse and remember content in a verbatim form. However, if you want to improve comprehension of the message, speaking the same content is the best way to improve memory. These findings support Craik and Lockhart's (1972) levels of processing theory in which content can be processed at a superficial level without a deeper understanding of the meaning of that content. It appears that singing, unless accompanied by enactive rehearsal such as what we did with *I'm a Little Teapot*, is a superficial learning technique while the use of language without singing is more likely to receive deeper processing. Thus, while singing provides scaffolds to verbal content, its effectiveness as an instructional feature depends on the kind of lesson to be learned. A challenge for instructional design is to create additional scaffolds between language and songs that can yield deeper processing of the content.

Parsing content: Sound effects and vocalizations as markers of important content. In addition to serving as a mode in which to represent content, formal production features can parse and mark content for further processing, thereby providing a scaffold to the verbal linguistic content that follows. Our early naturalistic study of *Fat Albert and the Cosby Kids* was the first place where we documented this beneficial effect (Calvert et al.,

1982). In the episode that we studied, Fat Albert would say, “Hey, Hey, Hey, I’ve got something to say.” Then he would say important verbal content that helped children understand the story. Children who selectively attended the most immediately after character vocalizations (in this case, Fat Albert’s “Hey, Hey, Hey”) understood the central story content the best. We hypothesized that children initially attend to vocalizations because of their perceptually salient qualities. Even after children were familiar with sound effects, we argued that these perceptually salient audio features still captured attention and elicited active processing because the sound effects had become a learned signal that was associated with important story content.

In subsequent experimental studies, we manipulated sound effects inserted at key scene transitions to see if it helped children understand the central plot-relevant content of the narrative. In one study, we found that sound effects inserted at three key scene changes increased kindergartners’, but not 3rd and 4th graders’, recognition of the implicit, central plot-relevant content (Calvert & Gersh, 1987). In a follow-up study in which we varied how rapidly the scenes and characters changed, we found that sound effects worked best for the rapidly-paced program (Calvert & Scott, 1989). Specifically, young children were more likely to selectively attend to scene changes in the rapidly-paced program when sound effects were present, not absent. Moreover, selective attention at these key program transitions predicted kindergartners’ comprehension of the content. Older children, by contrast, did not need the sound effects for comprehension of the rapidly-paced program. Recent research indicates that infants begin to integrate sound effects with target actions at about age one (Somander, Garcia, Miller & Barr, 2005). Taken together, the results suggest the value of sound effects as a way to draw attention

to key program content that can then lead to temporal integration of the plot line. Put another way, sound effects and character vocalizations can help children build a scaffold to link important program transitions and fill in the gaps when they are viewing television programs and use of these salient features to guide attention begins early in development.

Does Literacy Involve More than Reading and Writing Text in the Information Age?

Our traditions as a species are deeply rooted in our creation of and use of written language, an aspect of thinking that makes us unique from other species. Written words as a mode of cultural transmission became dominant when the printing press emerged, allowing that aspect of thought to be widely disseminated with ease (Calvert, 1999). We now live in a world where not only can we view other people's thoughts through production practices like flashbacks in time, but we can also create and transmit our visual realities through devices like cameras or drawings that we interface with computers. Put another way, newer technologies allow us to communicate with one another in visual as well as in musical and written forms of thought. There is much less known about how we come to understand and use non-verbal symbols, even though they comprise much of what children experience in their daily lives.

Take a feature like a camera zoom versus a camera dissolve. The camera zoom provides focus; it simulates how one uses one's eyes to learn information. A camera zoom moves into a close-up of an object, modeling the skill of whole to part or vice versa (Salomon, 1974). By contrast, camera cuts call upon viewers to fill in the gap of going from part to whole. Salomon's classic research (1974) demonstrated that young preschool aged children benefited the most after exposure to a camera zoom which supplanted (i.e., provided) the cognitive skill for them whereas older children benefited

the most from viewing a camera cut which activated that cognitive skill, calling upon children to produce the activity of going from part to whole themselves.

Camera dissolves are a media convention that generally represents a shift in time, including major time changes such as flashbacks. Although the dissolve does not have a direct link to how we use our eyes, it is a representation that is easier to understand than a camera cut when conveying a flashback. For instance, Calvert (1988) found that kindergartners and first graders were much more likely to understand that a flashback had taken place when a dreamy dissolve rather than an abrupt camera cut had occurred between the scene transitions. While fourth and fifth graders were less dependent on the camera cut for comprehension of the time change, even they understood the camera dissolve best. How do children come to understand camera dissolves or camera cuts? We don't really know.

Implications for Instructional Design

As can be seen from the empirical research, production features provide a vast array of options to provide scaffolds for motor, verbal, visual, and sequential ways of learning. Pauses and prompts that are built into a program can elicit active processing from very young children who interact with the characters and content, thereby yielding enactive ways of remembering content and imitative displays of that learning. Action provides a visual mode to represent content that can facilitate but also disrupt memory of linguistic information, depending on how those scaffolds are built between these two symbol systems. Singing provides a reflective way of processing words, though it can stay superficial unless enactive bridges are built to elicit deeper processing of the content. Sound effects and character vocalizations elicit attention and facilitate processing of

central and sequentially presented content though they can be distracting if they do not match the ecology of the visual presentation.

One interesting design challenge is how to facilitate children's attention to, and learning of, written rather than spoken words on a screen. I've observed poorly designed interfaces where the action is being presented at the same time as the written words are appearing on screen. The goal of the design is to get children to read, but the presentation is probably drawing attention away from the words and to the moving images. The solution may well be to have the words appear while a still image is on screen and then have the images move after the words are read and written. If the words are appearing as they are being read, the moving words should elicit attention and processing, thereby improving word learning. Faces are also distracting as children tend to look at the person rather than the words on the screen. Electric Company dealt with this design issue by having two faces appear sideways and in shadow, thereby facilitating attention to written words that were blended on the screen rather than to the faces.

As an instructional design feature, singing has the educational benefit of preserving a verbatim memory of events that can be accessed for very long periods of time. The design challenge of singing is to get children and even adults to think more deeply about the content so that they not only remember the lyrics, but understand the message. Fortunately, words can be processed at any time so that information is available if listeners are prompted or motivated to think about the underlying message. At young ages, enacting the lyrics helps children get the right message (Calvert & Goodman, 1999), and repeating the song helps enhance verbatim memory of the content (Calvert & Tart, 1993). While we have not studied singing that is accompanied by the

written words, that too may be helpful for yielding deeper processing at older ages. Indeed, sometimes it is difficult to understand the lyrics of songs unless you can see a verbal written display of the lyrics.

As instructional design features, singing and vocalizations can readily elicit children's visual attention, thereby providing a bridge to the significant program content that improves learning. Perceptually salient sounds are a better ecological fit when embedded within rapidly-paced television programs (Calvert & Scott, 1989). They are most useful at very young ages in terms of plot comprehension, but they elicit attention and processing at many developmental time frames, starting in infancy and continuing to attract attention during middle childhood (Somander et al., 2005; Calvert et al., 1982). Sound effects are easily integrated into previous productions, which we often did during our studies, making them a very cost effective way to improve the instructional effectiveness of already existing programs.

Conclusion

Children spend much of their daily lives with screen media. At this point, television is still the dominant medium in children's homes, particularly those from African American and Latino families (Roberts et al., 2005), and children from these ethnic minority groups tend to be less successful in school than their Caucasian peers. Much has been done via programs like Sesame Street to improve the educational outcomes of low-income children (Fisch & Truglio, 2001), in part through the judicious use of production features that attract children's attention and that provide scaffolds to link different forms of thought that dominate at different points in development. While the future promises to be a more interactive one, our research indicates that similar design

principles apply to computer as well as television platforms. Features can be used to elicit active participation, to provide visual and verbal modes to represent content, and to parse the content in ways that attract attention and processing of important linguistic content.

In the 21st century, dominated by visual media, it is timely to think about how our newer technologies are influencing literacy, both in its traditional verbal form as well as in its visual and increasingly interactive form. The judicious use of formal features as scaffolds to the content and to the form of children's thought are essential to the creation of well-designed instructional platforms and to our understanding of what literacy entails in the 21st century.

References

- Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R. & Hill, J. (2002). The exact mind: Empathizing and systematizing in autism spectrum conditions. In U. Goswami (Ed.). *Blackwell Handbook of Childhood Cognitive Development* (pp. 491-508). Malden, MA: Blackwell.
- Berlyne, D.E. (1960). *Conflict, curiosity, and arousal*. New York: McGraw Hill.
- Bruner, J.S., Olver, R.R. & Greenfield, P.M. (1968). *Studies in cognitive growth*. New York: Wiley.
- Calvert, S.L. (1991). Presentational features for young children's production and recall of information. *Journal of Applied Developmental Psychology, 12*, 367-378.
- Calvert, S.L. (1994). Developmental differences in children's production and recall of information as a function of computer presentational features. *Journal of Educational Computing Research, 10*, 131-143.
- Calvert, S.L. (1999). *Children's journeys through the information age*. Boston: McGraw Hill.
- Calvert, S.L. (2001). Impact of televised songs on children's and young adults' memory of educational content. *Media Psychology, 3*, 325-342.
- Calvert, S.L. & Billingsley, R.L. (1998). Young children's recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology, 19*, 97-260.108.
- Calvert, S.L. & Gersh, T.L. (1987). The selective use of sound effects and visual inserts for children's television story comprehension. *Journal of Applied Developmental Psychology, 8*, 363-375.

- Calvert, S.L. & Goodman, T. (1999, April). Enactive rehearsal for young children's comprehension of songs. Poster presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, New Mexico.
- Calvert, S.L., Huston, A.C., Watkins, B.A. & Wright, J.C. (1982). The relation between selective attention to television forms and children's comprehension of content. *Child Development, 53*, 601-610.
- Calvert, S.L. & Scott, M.C. (1989). Sound effects for children's comprehension of fast-paced television content. *Journal of Broadcasting and Electronic Media, 33*, 233-246.
- Calvert, S.L., Strong, B.L., Jacobs, E.L. & Conger, E.E. (in press). Interaction and participation for young Hispanic and Caucasian children's learning of media content. *Media Psychology*.
- Calvert, S.L. & Tart, M. (1993). Song versus prose forms for student's very long-term, long-term, and short-term verbatim recall. *Journal of Applied Developmental Psychology, 14*, 245-260.
- Calvert, S.L., Watson, J.A., Brinkley, V. & Bordeaux, B. (1989). Computer presentational features for young children's preferential selection and recall of information. *Journal of Educational Computing Research, 5*, 35-49.
- Calvert, S.L., Watson, J.A., Brinkley, V. & Penny, J. (1990). Computer presentational features for poor readers' recall of information. *Journal of Educational Computing Research, 6*, 287-298.
- Craik, F. & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 521-533.

- Fisch, S. & Truglio, R. (2001) *G is for Growing: Thirty years of research on children and Sesame Street*. Mahwah, NJ: Lawrence Erlbaum.
- Gibbons, J., Anderson, D.R., Smith, R., Field, D.E. & Fischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual material. *Child Development, 57*, 1014-1023.
- Hayes, D. & Birnbaum, D. (1980). Preschoolers' retention of televised events: Is a picture worth a thousand words? *Developmental Psychology, 16*, 410-416.
- Huston, A.C. & Wright, J.C. (1983). The informative functions of formal features. In J. Bryant & D.R. Anderson (Eds). *Children's understanding of television: Research on attention and comprehension* (pp.35- 68). New York: Academic Press.
- Lawler, R.W. (1982). Designing computer-based microworlds. *Byte*, 138-160.
- Meltzoff, A.N. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. *Blackwell handbook of childhood cognitive development* (pp. 6-25). Oxford: Blackwell.
- Meltzoff, A.N. (1988). Imitation of televised models by infants. *Child Development, 59*, 1221-1229.
- Moore, M. & Calvert, S.L. (2000). Vocabulary acquisition for children with autism: Teacher or computer instruction. *Journal of Autism and Developmental Disorders, 30*, 359-362.
- Neuman, S. B., Copple, C., & Bredekamp, S. (2000). *Learning to read and write: Developmentally appropriate practices for young children*. Washington DC: National Association for the Education of Young Children.

- Roberts, D., Foehr, & Rideout, V. (2005, March). Generation M: Media in the lives of 8-18 year-olds. The Henry J. Kaiser Foundation.
- Salomon, G. (1974). Internalization of filmic schematic operations in interaction with learners' aptitudes. *Journal of Educational Psychology*, 66, 499-511.
- Somanader, M., Garcia, A., Miller, N., & Barr, R. (2005). *Effects of sound effects and music on imitation from television during infancy*. Paper presented at the Society for Research on Child Development, Atlanta, GA.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Table 1.

A Taxonomy of Formal Features

FEATURES	DEFINITIONS
Macro Features	
Action	amount of movement
Pace	rate of scene and character change
Visual Micro Features	
Cuts	camera technique involving quick shifts in visual perspective
Zooms	continuous camera technique moving the lens toward or away from an object or scene
Fades	camera technique that goes to black
Dissolves	camera technique that makes the edges of the scene blurry as a new image emerges
Pans & trucks	camera technique where objects or events are followed continuously on a horizontal or vertical plane
Visual Special Effects	camera techniques such as trick photography, freeze frames, and fast motion
Auditory Micro Features	
Foreground Music	loud music with no dialogue
Background Music	music with dialogue
Vocalizations	non-speech noises
Sound Effects	unusual prominent audio effects (e.g., zip!)

Laugh Track	audible laughing by an off-screen audience
Child Dialogue	child speech
Adult Dialogue	adult speech
Non-human Dialogue	speech by an animal character or other non-human character
Narration	speech by one individual, typically explaining on-screen events
Singing	music and lyrics combined