International Journal of Humanities and Social Science Review; Vol. 9 No. 1; August 2023
ISSN 2415-1157 (Online) ISSN 2415-1335 (Print)
Published by Research Institute for Progression of Knowledge

An Information Processing-Based Theoretical Framework for Studying Educational Television and Emotion-Recognition Capability of Children with Autism Spectrum Disorder

Erin L. Ryan, PhD
Kennesaw State University
1000 Chastain Road
Kennesaw, Georgia 30144
USA

Abstract

Baron-Cohen (2008) and colleagues suggest children 2-8 with Autism Spectrum Disorder (ASD) can learn skills like emotion recognition from media content like Thomas the Tank Engine. Whereas these researchers posited in their Hyper-Systemizing Theory that the reason for learning lies in idiosyncrasies related to ASD, this author contends that the manner in which the media is crafted is the key to success. To test this notion experimentally, a "hybrid theory" specific to the learning needs of children with ASD is necessary. This paper thus fuses elements from the following information processing theories with Hyper-Systemizing Theory to create such a hybrid: Cognitive Load Theory, the Limited Capacity Model of Motivated Mediated Message Processing, Capacity Model, and the Cognitive Theory of Multimedia Learning. The goal is to craft a theoretical model to serve as the basis for future work in the area of ASD and learning from mediated content.

***Keywords*:** Autism, ASD, emotions, educational television, learning, information processing

An Information Processing-Based Theoretical Framework for Studying Educational Television and Emotion-Recognition Capability of Children with Autism Spectrum Disorder

1. Introduction

How do young children on the Autism Spectrum respond to media messages? Promising research out of the United Kingdom features an expert on Autism Spectrum Disorder (ASD) who uses the children's educational program *Thomas the Tank Engine* as inspiration for his research on teaching children Executive Functioning (EF) skills like mind-reading, empathy, and perspective-taking (Baron-Cohen, 2008). This sparks the question: Is it really possible that programs like *Thomas The Tank Engine* can actually *teach* children with ASD to develop their EF skills? A search for scholarly literature uncovered countless articles in both the popular press and academic journals written by the same ASD expert from the UK, but curiously nothing written by communication or media scholars. First stop: a useful theoretical framework.

This task would require becoming familiar with the ASD research first, followed by the EF research, and then an investigation of appropriate media research to tie it all together. Much research in this vein uses Bandura's Social Learning Theory work as a basic model, but what are the optimal conditions under which that kind of observational learning can occur, given this special population? The hope was that an existing media theory would work as an appropriate framework upon which to build the project. However, none of the theories, as defined by media scholars, quite fit the bill. Below is a brief outline of the research in all three areas which lead to the conclusion that no existing theory could, by itself, frame such a study. A "hybrid" theory was necessary.

2. The Autism Spectrum and Executive Functioning Deficits

Focusing on this special population of children first necessitates an understanding of ASD and how these conditions often result in difficulties with EF, and how that then impacts their relationship with media. For that, the author first turned to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5).

In 2013, after a 14-year revision process, the American Psychiatric Association (APA) updated their DSM-IV definition of ASD from a four-disorder diagnosis (autistic disorder, Asperger's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified) to a single "umbrella" disorder simply referred to as ASD in the new DSM-V (APA, 2013). Researchers at the National Institute of Mental Health found that the four diagnoses were not consistently applied across treatment centers and believed that the single umbrella disorder would lead to improvements in the proper diagnosis of ASD. Individuals diagnosed as such will still fall on a continuum, with some children displaying mild symptoms and others more severe.

According to the APA (2013), children diagnosed with ASD display several communication deficits. Many times these children respond inappropriately in conversations and they tend to misread (or fail to read) nonverbal signals, which greatly impairs their ability to build friendships with other children and effectively communicate with adults. Haney (2013) explained that regardless of where a child falls on the ASD spectrum, she displays three core deficits including difficulty with socialization, deficits in communication (verbal and nonverbal), and "restrictive, repetitive, and stereotyped patterns of behaviors and interests" (p. 39).

Regarding socialization, one of the earliest indicators of ASD is atypical development in early childhood with joint attention, social referencing, and symbolic and pretend play, three key EF skills. Dawson et al. (2004) explained that joint attention involves gazing and pointing and typically begins in infancy; it is the basis of engaging in shared experiences. Social referencing involves looking to others' emotions before acting or to gather information about current events (Dickstein & Park, 1988). Difficulty with joint attention, social referencing, and the engagement with pretend play diminishes a child's ability to engage in social reciprocity, often resulting in rejection and alienation from non-ASD peers (Haney, 2013). "Theory of mind," (TOM) sometimes called "mindblindedness" is also a core deficit where socialization is concerned. TOM refers to the ability to understand the mental state of someone else and virtually "engage in a form of mind reading" (Durand, 2005, p. 92). When this ability is lacking, social relationships are difficult to maintain.

Deficits in overt communication also hinder children with ASD. Regardless of overall cognitive ability, these children have difficulty maintaining eye contact, accurately using and regulating nonverbals such as facial expressions, postures, and gestures during social interactions, and struggle with reading the body language of others. They also tend to have difficulty with "social and emotional reciprocity" and often fail to appropriately navigate the give-and-take of social relationships (Haney, 2013, p. 42). Children with ASD often lag behind their peers with the onset of spoken language as well, which results in difficulty initiating and sustaining conversations (Christopher, Sears, Williams, Oliver, & Hersh, 2004). Haney (2013) also noted that children with ASD may have an unusual quality to their speech, such as unusual pitch and volume or a mechanical quality, and their language may be repetitive.

Lastly, children with ASD can be overly reliant on routine, become intensely focused on minutiae or inappropriate items, and are highly sensitive to any change to their immediate environment (APA, 2013). Self-stimulatory repetitive behaviors such as hand or finger-flapping or body rocking are quite common and many children with ASD prefer objects to people and display restrictive interests that are atypically intense in focus (such as only playing with trains, talking about trains, drawing trains, etc. for several months) (Heflin & Aliamo, 2007). Rigid adherence to rituals or routines can be particularly challenging for caretakers of children with ASD, as these children often display discomfort and anxiety when these routines are disrupted in some way (Gabriels, Cuccaro, Hill, Ivers, & Goldson, 2005).

3. Learning from Educational Media by Children with ASD

Underscoring the media side of this project is Bandura's (1977) Social Learning Theory framework. Countless studies of non-ASD children lend support for the notion that children can learn through observation, and they demonstrate such knowledge via imitation/ modeling behavior. The use of video modeling is a well-validated behavioral intervention and has been successful in teaching social skills to children with non-ASD social deficits (e.g., Dorwick & Jesdale, 1991).

But video modeling also appears to be beneficial for teaching skills to children with ASD. Research has documented the successful acquisition of vocalization and communication skills, (Charlop & Walsh, 1986; Charlop & Milstein, 1989), social and play skills (D'Ateno, Mangiapanello, & Taylor, 2003; Taylor, Levin, & Jasper, 1999; Wert & Neisworth, 2003), academics (Kinney, Vedora, & Stromer, 2003) and adaptive behavior (Shipley-Benamou, Lutzger, & Taubman, 2002) by children with ASD. Perhaps most crucial to the Thomas Project, however, are the studies pointing to successful imitative learning from video of emotion processing (Corbett, 2003) and perspective taking (Charlop, Christy & Daneshvar, 2002; LeBlanc, Coates, Daneshvar, Charlop-Christy, Morris, & Lancaster, 2003) by children with ASD. These videos were specifically crafted to be instructional, however, whereas *Thomas* is animated entertainment first and foremost. So the question remains: can we be confident that entertainment programming not specifically crafted for an ASD audience will provide appropriate video models to encourage the learning of EF skills like emotion-reading and Theory of Mind? As these skills are crucial for communication and the maintenance of social relationships, and very often *Thomas* can be found on television on PBS or on video at local libraries (both free resources for parents), it is important to craft an appropriate theoretical framework for experimental study.

4. Hyper-systemizing Theory

It seems fitting to begin with the work of UK Autism researcher Simon Baron-Cohen when searching for a theoretical model to study this phenomenon. Whereas Baron-Cohen's work on hyper-systemizing is not media-centric in its approach, his ideas about the core difficulties experienced by children with ASD must be a part of the conversation. Hyper-systemizing theory suggests that individuals with ASD have an incredibly strong drive to systemize. Baron-Cohen (2008) defines a system as "something that takes inputs and delivers outputs" (p. 65). Through a system, and learned rules of cause and effect, the human brain observes an input and makes a prediction of the potential output based on probability (Baron-Cohen, 2008). These observations lead to the identification of laws. Baron-Cohen (2008) explained, "Systems that are 100% lawful have zero variance, or only 1 degree of freedom, and can therefore be predicted (and controlled) 100%" (p. 66). When the variance is wider, as in a social interaction, there is an increase in degrees of freedom and a decrease in predictability. Lawful predictably is significant to individuals with ASD because of their "need for sameness" or "resistance to change" (Kanner, 1943).

The hyper-systemizing theory proposes that every human brain has a systemizing mechanism (SM) (Baron-Cohen et al., 2005). Each individual, dependent upon mental ability, has this mechanism set at a different "level" ranging from level 1, which is characterized by little or no desire to systemize, to level 8, which shows a strong affinity for lawful, systematic change and an intolerance to hasty, unpredictable change (Baron-Cohen et al., 2005). Depending on the severity of one's autism, an ASD individual will operate on a relatively high systemizing level (6-8), while most neurotypical people have a limited attraction to lawful systemizing and operate on level 2 or 3. Furthermore, the higher the SM level, the lower one's ability to generalize (Baron-Cohen et al., 2005). Simply stated, the hyper-systemizing theory is meant to measure the effect of different kinds of change (lawful, unlawful, or somewhere in between) on individuals of different mental capacities on and off the autism spectrum. In addition, there are two types of change: agentive and non-agentive. Baron-Cohen et al. (2005) explained, "If an object change is perceived to be self-propelled, the brain interprets the object as an agent, with a *goal*" (p. 3). Agentive change is interpreted by the human mind via special neurocognitive mechanisms, commonly known as the "empathizing system" (p. 4). In contrast, a non-agentive change is not self-propelled, but uses systemizing processes to search for structure (Baron-Cohen et al., 2005).

The actual process of systemizing involves five phases: analysis, operation, repetition, law derivation, and confirmation/disconfirmation (Baron-Cohen et al., 2005). Phase 1, analysis, focuses on observing inputs and outputs (instruments that create change). Phase 2, operation, centers on performing an operation on the input and observing the change on the output. Phase 3, repetition, performs the same operation multiple times to test the established pattern of the input and output. Phase 4, dictated law derivation, references the law derived from the observed change (i.e., if X, then Y). The final phase, confirmation/disconfirmation, suggests that "if the same pattern of input-operation holds true for all instances, the law is retained" (Baron-Cohen et al., 2005, p. 5). If the pattern is different, phases 2-5 are repeated, and the law is revised.

So why is hyper-systemizing important to this project? Baron-Cohen and his colleagues have used the knowledge that preschool children with ASD are hyper-systemizers to create a series of DVDs called "The Transporters" to teach emotion recognition (Baron-Cohen, Golan, Chapman, & Granader, 2007). The logic is that children with ASD tend to avoid looking at the human face, and thus find it difficult to understand why faces "move" the way they do. The inability to read others' emotions impairs successful socialization. But since researchers know that these hyper-systemizers are looking for "rules" and patterns first and foremost, the characters in the DVDs play off of their fascination with things like rotating wheels, spinning tops, rotating fans, and mechanical, lawful motion and include vehicles that run on tracks or lines (such as trains). Each of these toy characters has an animated human face, and the episodes focus on 15 human emotions: happy, sad, angry, afraid, excited, disgusted, surprised, tired, unfriendly, kind, sorry, proud, jealous, joking, and ashamed. Baron-Cohen et al. (2007) reported that after four weeks of watching the videos 15 minutes per day, children with ASD ages two to eight caught up with typically-developing children of the same age in their performance on emotion-recognition tasks. This has important implications for the viability of *Thomas* to do the same. However, Baron-Cohen and colleagues explain their results only in terms of hyper-systemizing. The current author agrees that the systemizing aspect most assuredly plays a role, but believe that there is more to the process. The style in which the material is presented plays a crucial role as well.

5. Cognitive Load Theory

To move beyond the systemizing explanation, the author looked to the media and communication literature for theories that seek to provide a more media-based explanation for why some material is more accessible to children with ASD. First up was Cognitive Load Theory (CLT).

CLT is an instructional theory based on the architecture of human cognition (Sweller, 1988; Leahy & Sweller, 2011). The theory borrows from aspects of information processing theories and, in a nutshell, emphasizes the inherent limitation of a person's working memory load on learning during instruction. Paas, Renkl, and Sweller (2004) explained that during complex learning, the amount of information that is processed can either "underload" or "overload" the amount of working memory available to a person. If people can build upon what they already know (called schemas), they can learn more effectively. However, "cold" learning is more taxing. Sweller's research on problem solving has indicated that the most common strategy used in problem solving is a strategy called a means-ends analysis, which requires a large amount of cognitive processing capacity because it does not employ schemas. Thus, Sweller suggested that anyone designing instructional materials can prevent this unnecessary cognitive load by *not* involving problem solving in the learning process. That is, make it as easy as possible for the learner to devote all of their working memory to the actual skill you are trying to teach, without distraction (Paas et al., 2004).

Through the use of CLT, media researchers are beginning to understand how learning with animation occurs and how best to optimize performance. Animation may actually *impede* learning considering "animations are often cognitively very demanding, resulting in decreased learning outcomes" (de Koning et al., 2007, p. 731). Quantitative data supports the use of animation for learning with the addition of attention cues to improve comprehension and test performance, but in order to do so, one must understand the theoretical basis of CLT and sources of cognitive load.

There are three sources of cognitive load involved with animation-based learning (de Koning et. al, 2007; Hoffler & Leutner, 2007). The first, intrinsic cognitive load focuses on the interaction of different learning elements. Secondly, extraneous cognitive load results from additional, unnecessary cognition use that is not directly related to learning the concept at hand, such as highly populated visuals and extraneous audio (Hoffler & Leutner, 2007). Collectively, intrinsic cognitive load and extrinsic cognitive load significantly hinder animation-based learning by including unnecessary visual information and other elements utilizing cognitive resources that could otherwise be focused and employed toward learning the concept. Lastly, germane cognitive load refers to learning knowledge in an instructional format that enhances information comprehension, while limiting supplementary information that needlessly occupies much needed cognitive resources (Hoffler & Leutner, 2007).

With the CLT framework in mind, animations designed in accordance with germane cognitive load are less demanding of limited cognitive resources and minimize the occurrence of extraneous distractors. The most empirically supported method to achieve and maintain the learner's attention is by application of the attention-guiding principle, which aids the learner through the use of cues. Bjorn de Koning et al. (2007) explained, "By adding a visual cue to a complex animation visual search should be reduced, thereby reducing extraneous load and allowing more cognitive resources to be allocated to learning" (p. 732). Given CLT then, children with ASD have the best chance of learning from *Thomas* if it is designed in accordance with germane cognitive load.

6. The Limited Capacity Model of Motivated Mediated Message Processing

Although it was not initially intended to be applied to children, Lang's limited capacity model has also provided a significant foundation for understanding the relationship between children and educational media content. Many of the concepts mirror CLT, but the Limited Capacity Model of Motivated Mediated Message Processing (LC4MP) explains the information-processing element a little differently. The preliminary concepts behind this model were introduced in the late 1970s (i.e., Shriffin & Schneider, 1977), but were not thoroughly examined until the early 1990s (i.e., Lang, 1992).

The resulting LC4MP model holds several basic assumptions. Firstly, people are information processors with a limited number of cognitive resources available to process information they receive (which is similar to the tenets of CLT), and the actual processing of information functions through three, dynamic sub-processes: encoding, storage, and retrieval; and two mechanisms: orienting behavior and resource allocation (Basil, 1994; Lang, 2000; Lang, 2006). Lang (2000) explained, "The orienting response is made up of an organized set of behavioral and physiological responses. Physiological measures can be used to measure the occurrence of an orienting response during television viewing" (p. 55). Resource allocation posits that processing information requires available resources within the individual to be distributed among the sub-processes for message interpretation (Lang, 2000). The model suggested that the three sub-processes simultaneously compete for an individual's fixed pool of total mental resources, and the application of these sub-processes are constantly changing as human behavior, communication, attention, and motivation continually evolve (Lang, 2006).

6.1 Encoding

Encoding is the process of selecting information from the environment and bringing it into the brain for further processing (Lang, 2000). Information is selected from a person's environment, and important aspects of that information are encoded into the person's knowledge base (Lang, 2006). This is an incomplete, selective, and ongoing process, as people continually select new information for processing. The model suggests that encoding is divided into three stages. First, the sensory receptors are engaged by the message (Eysneck, 1993). Lang (2000) explained, "This can be thought of as exposure or perception. Information gathered by the sensory receptors enters some kind of sensory store" (p. 48). Each sense has a virtually unlimited amount of sensory stores that hold more information than a person can process, meaning that "only a fraction of the information held in the sensory stores moves on into active (or short-term or working) memory" (Lang, 2000, p. 48). As individuals constantly receive new information, bits that are not selected for further processing are soon replaced by new incoming data (Lang, 2000). In the final stage, the sensory store's selected information is further processed and enters short-term "working" memory. In order for the information to be encoded, mental resources must be allocated to process it. Such information can be selected for encoding through response. An orienting response, or individual's reaction to a stimulus, consists of behavioral, cognitive, and physiological changes in the receiver (Lang, 2006; Lang & Basil, 1998).

6.2 Storage

Lang (2000) explained, "The more a person links a new bit of information into this associative memory network, the better that information is stored. This process of linking newly encoded information to previously encoded information (or memories) is called storage" (p.50). Bits of encoded information are merely mental representations of the information received (Lang, 2006). In the storage process, these newly encoded representations are linked with previously encoded, older information. Links are created between the new and old information which results in a network of stored information. So, the more links associated with the new information, the more thoroughly the information is stored in long-term memory (Lang, 2000). Also, unless intentionally allocated, few resources are devoted to this storage process often causing many bits of information to be stored poorly (Lang, 2006).

6.3 Retrieval

The third sub-process, retrieval, is marked by the recovery of previously stored mental representations from the memory network (Lang, 2006). In order for the reactivation of the memory to occur, resources must once again be allocated to process it. When the allocation occurs, and the specific memory is reactivated, the activation spreads from the information bit to related linked information. It is much easier to retrieve information that has been more thoroughly stored with a greater number of links in the memory network (Lang, 2006).

There are two types of retrieval that can occur when watching a mediated message: “later retrieval” of message content and “concurrent retrieval” of already-known information (Lang, 2000, p. 54). Later retrieval is not constrained by time or resources, nor does it have any influence on the encoding or storage stages as it does not occur until after the message is viewed. However, concurrent retrieval perpetually retrieves information during viewing, which directly influences storage and encoding (Lang, 2000). An individual’s goal and the message content and structure are just a few of the factors that heavily influence automatic and controlled allocation and reallocation of resources (Lang, 2006).

Lang’s work on LC4MP appears to be an evolution of the CLT and its tenets. Both acknowledge that people are information processors with a limited amount of cognitive resources readily available to process new information. But whereas CLT focused on the use of schemas, the LC4MP focuses on the encoding, storage, and retrieval of new information and how these processes compete for a person’s finite amount of mental resources. As a person encodes and links new information to previously-learned material, it can be stored and later retrieved; the more links, the better. Providing these links appears to be key when using media to “teach” children. Particularly for children with ASD, fewer distractions that could compete with these links seems ideal when designing anything instructional. If the content in *Thomas* that is designed to teach emotion-recognition is simple and repetitive, according to the LC4MP it should be successful.

7. Fisch’s Capacity Model

Contributing to the conversation, Fisch’s Capacity Model also focuses on information processing but is the only theoretical model specific to children’s learning from mediated content. In addition to examining the information-processing element, the Capacity Model seeks to explain how children actually extract and comprehend educational media messages. Fisch (2004) also placed additional emphasis on the message itself and the subsequent learning of different types of content.

Much like both CLT and the LC4MP, Fisch’s Capacity Model posits that working memory possesses a limited amount of resources that are available for allocation at any given time. Therefore, if the demand of a given task (or cognitive load) exceeds the working memory’s available resources, the material cannot be processed effectively (Fisch, 2004; Gopher & Donchin, 1986; Hart, 1986). According to the Capacity Model, a child’s comprehension depends on the cognitive demands of simultaneously processing both *narrative* and *educational* content (Fisch, 2004). Narrative content refers to the story presented to the viewer during the program, whereas the educational content is the instruction itself.

According to the Capacity Model, the allocation of working memory resources when processing narrative and educational content is influenced by three factors. First, narrative content is a default priority in working memory, as resources will almost always be allocated to processing it first. Second, when narrative dominance occurs, and there are high demands for processing the narrative, there are fewer resources available for processing educational content and the two must compete for resources (Fisch, 2004). However, if there is low demand for processing narrative content, (i.e., few transitions, a slow pace, etc.), then there are more resources available to allocate to the processing of educational content (Fisch, 2004). Finally, resources can be allocated for educational content voluntarily, depending on parental commentary, reasons for viewing, and comprehension of the material (Fisch, 2004).

Fisch's Capacity Model is comprised of three basic elements: the processing of narrative, the processing of educational content, and the distance between the two. The concept of distance is a unique component of the model and refers to "the degree to which the educational content is integral or tangential to the overall narrative in a program" (Fisch, 2004, p. 151). Distance can be described in terms of how educational content is causally related to the structure of the story (Fisch, 2004). Fisch (2004) identified two types of distance: small and large distance. A small distance exists when educational content is causally connected to a large number of subsequent events. A large distance occurs when educational content is embedded in the plot in a manner that does not forward the story (Fisch, 2004). A small distance (i.e., well-integrated content) minimizes the mental resources needed to comprehend educational content, making it easier for children to process and retain both aspects of the message (Fisch, 2004). According to the capacity model, if the distance between educational and narrative content is too large, then the two must compete for resources. In this case, working memory will more likely devote resources to processing the narrative, resulting in a decreased ability to process educational content (Fisch, 2004).

Furthermore, the Capacity Model identifies when comprehension of educational content will be the strongest. According to Fisch (2004), comprehension of educational content is stronger when: (1) the total amount of working memory resources devoted to understanding the material is increased; (2) demands for processing narrative content is small, leaving more resources available for educational content; (3) when the distance between narrative and educational content is small (4) when the viewer has a greater motivation to learn and voluntarily allocates more resources; and (5) when demands for processing educational content is small. Also, if the demands for processing the educational and narrative content exceed the capacity of working memory, then comprehension is impaired and the acquisition (encoding, storage, and retrieval) of the information decreases (Fisch, 2004).

7.1 Governing Principles

Fisch's Capacity Model is based on specific governing principles that influence the allocation of resources for processing of narrative and educational content. Principles such as narrative dominance, allocation of resources, prior knowledge, and inferences can significantly affect resource allocation (Fisch, 2004). First, the model posits that priority is given to the comprehension of narrative content, a principle called "narrative dominance" (Fisch, 2004, p. 153). This notion is based on the finding that children's comprehension of television is influenced by the motives for viewing (Salomon & Leigh, 1984). The allocation of resources, another governing principle, is based on the assumption that "viewers can choose to allocate resources differently among the processing of narrative and educational content" (Fisch, 2004, p. 155). The decision to allocate resources is based on a variety of factors that influence how children attend to the narrative. These factors include the motivation for watching, comprehension, parental commentary and co-viewing, among others (Collins, Sobol, & Westby, 1981, Kwaitek & Watkins, 1981; Salomon & Leigh, 1984). Therefore, depending on the circumstances, a child may be more motivated to allocate additional resources to educational content, yet the narrative content is not forgotten (Fisch, 2004). Also, the model places great emphasis on narrative dominance, so it is logical to assume that deficiencies in resource availability will most likely occur in processing educational content (Fisch, 2004).

In addition, a child's age, mental development, and cognitive capacity can greatly affect their information processing abilities (Fisch, 2004). As children age, they develop a greater ability to understand the narrative and educational content in a program and can hold increasing amounts of information in working memory (Dempster, 1981; Gathercole & Baddeley, 1993). This increase may allow for deeper processing of narrative and educational content, as well as a greater efficiency and effectiveness in managing these parallel processes (Fisch, 2009). Additionally, as children age they develop the ability to manage multiple tasks in working memory (Lawson & Kirby (1981).

The Capacity Model adds a new dimension to the information-processing element common to CLT and Lang's LC4MP by focusing on the content itself. The inclusion of the "distance" element where narrative and educational content overlap is unique to this theoretical model, and certainly informs the discussion regarding the possibility of *Thomas* to teach emotion recognition via an educational television program.

8. Cognitive Theory of Multimedia Learning

The focus on content that the Capacity Model offers is further fleshed out in the cognitive theory of multimedia learning (CTML). CTML was developed with the intent to help researchers and academics better understand the human brain's ability to learn via multimedia content. Through in-depth research, CTML was constructed by a group of cognitive psychology researchers, most prominently Dr. Richard Mayer of the University of California (Sorden, 2012). Mayer and his fellow academics founded CTML on the multimedia principle, which determined that "people learn more deeply from words and pictures than from words alone" (Mayer, 2001, p. 31). Mayer (2001) claimed that the current sophisticated multimedia technology provides an excellent format for structuring multimedia instruction so that it is most beneficial to the learner. However, the process for enhanced learning is more complex than simply pairing words and images, the content must be meaningful (Mayer, 2001).

When Mayer (2001) began his extensive research into CTML, he hypothesized that "multimedia instructional messages that are designed in light of how the human mind works are more likely to lead to meaningful learning than those that are not" (p. 46). He defined a multimedia instructional message as "a communication containing words and pictures intended to foster learning" (Mayer, 2001, p. 47). As Mayer and his colleagues continued to build the CTML, they established three essential assumptions.

8.1 Assumptions of CTML

In agreement with previous research (i.e., Mayer & Moreno, 1998; Mayer, 2001) CTML holds three underlying assumptions: the limited capacity assumption, the active processing assumption, and the dual-channel assumption. The dual-channel assumption posited that working memory consists of both visual and auditory channels that process separately, which relates to the theory of working memory established by Alan Baddeley (1986). Secondly, the CTML is based on the assumption of limited capacity (as in LC4MP), which can be understood as the minimization of the amount of information the brain can simultaneously comprehend. Finally, the assumption of active processing (as in CLT), which establishes that individuals comprehend new information in meaningful ways if they can remain attentive, organize the new information in a systematic and intelligible manner, and suffuse it with pre-existing knowledge (Mayer & Moreno, 1998).

As all of the theoretical models outlined here have posited, humans can only process a finite amount of information in a channel at one time, and we make sense of incoming information by actively creating mental representations. Mayer and Moreno (1998) focus on three memory stores: sensory memory, working memory, and long-term memory. Sensory memory is akin to short-term memory, as it receives stimuli and stores it for only a short time. His concept of working memory is quite similar to CLT and LC4MP in that this is where we actively process information to create mental constructs or schemas. And our long-term memory is if course where we store everything we've learned.

8.2 Five Stage Process

For Mayer and Moreno (1998), CTML captures the idea that our brains do not interpret a multimedia presentation of words, pictures, and sound in a mutually exclusive manner; rather, we continually select elements of the presentation to organize in such a way as to produce logical mental constructs. Mayer and Moreno (1998) underscore the importance of being able to integrate new information with prior knowledge, and believes this to be the best chance for learning, much like Lang's notion of linking in her LC4MP.

The actual method of CTML consists of five separate and distinct cognitive processes (Mayer & Moreno, 1998). The first requires the learners to utilize words relevant to the content in the working memory's verbal process. Second, one must select relevant, visual content to be processed by the working memory. Third, one systematically organizes the new information into a comprehensible and memorable verbal model. Fourth, one methodically organizes selected visuals into a "pictorial model." Lastly, one must undergo a holistic amalgamation of one's pre-existing knowledge, and the newly established meaningful visual and verbal (mental) representations (Mayer & Moreno, 1998). To aid this learning process, Mayer and Moreno's (1998) advice to content creators is to provide coherent verbal, pictorial information that guides the viewers to select relevant words and images, thus reducing the cognitive load for any one processing channel. So, if the creators of *Thomas* have crafted their program with meaningful visual and verbal representations that child viewers can link to their existing notions of emotion recognition, the CTML would predict successful learning for children with ASD.

9. Conclusions

Clearly there are many facets to the theoretical understanding of the information processing that occurs in the minds of children with ASD when they attempt to learn from mediated content. Baron-Cohen's work on hyper-systemizing tells us that the mere fact that Thomas (and most other characters) is a train with spinning wheels and "rules" (i.e., stays on a track, can only go two directions) makes it likely that children who are high systemizers will pay attention to and relate to this character. And children with ASD who have difficulty maintaining social relationships (i.e., due to lack of eye contact, inability to read nonverbals, language delays, etc.) should be more likely to learn from such a character rather than from a live model. But this only tells half of the story.

What is it about the content itself and how it is presented that appears to make programs like *Thomas* such a success in teaching emotion recognition to children with ASD? CLT, LC4MP, Capacity Model, and CTML each offer explanations that content creators can utilize to achieve such a goal. Keeping in mind the limitations of this special population in terms of EF skills like mindreading and perspective-taking, it appears that an educational program's best chance of success is to keep things simple so that viewers have the cognitive resources to link the content to previous schemas, as CLT, LC4MP, and CTML would suggest. Additionally, the Capacity Model's focus on the notion of narrative distance is important to consider. A child viewer's best chance to learn from educational content is to ensure that the instructional and narrative content are intertwined so as to make it as easy as possible for viewers to learn the instructional material. The distance element is crucial.

In fact, Hoffler and Leutner's (2007) research on static pictures versus animation found a clear advantage for "representational" animations (i.e., when the topic to be learned is explicitly depicted, rather than when the animation serves only a decorational function) over static pictures when it came to instructional effectiveness. This also speaks to CTML's focus on audio vs. video channels and the importance of aligning the two so they work together to "teach" new material. Representational animations facilitate the generation of a mental model of the topic to be learned by providing a prototype (Hoffler & Leutner, 2007). If the animation is used solely for decorational purposes, where the video is unrelated to the audio instruction, it becomes more of a distraction. Following the tenets of CLT, the decorational animation would impose "extraneous cognitive load" on the viewer, unnecessarily burdening working memory capacity (p. 734). Mayer's (2005) work with CTML certainly shows strong empirical evidence that instructional outcomes are greatly improved by coordinating the presentation of pictorial and verbal information, what he terms the "multimedia principle."

So which pieces of these theories are best integrated to form a workable "hybrid theory of cognitive capacity and multimedia learning by high-systemizers" for studying educational children's programming like *Thomas* and ASD? All four focus on working memory and the limitations of the cognitions of the viewers and there are definitely elements of each that form an impressive framework when combined. First, CLT's focus on schemas and "germane cognitive load" reminds content creators to limit extraneous or supplementary information when crafting educational media content. Second, the LC4MP outlines the encoding, storage, and retrieval processes that continuously compete for a viewer's fixed pool of total mental resources, and highlights the importance of forming links between existing knowledge and new material. Third, the Capacity Model's contribution of "narrative distance" reminds researchers and content creators how crucial it is that the narrative and instructional content be closely intertwined if child viewers have any chance of learning the educational lessons presented. And lastly, the CTML's "multimedia principle" posits that viewers learn more deeply from words *and* pictures than from words alone.

Keeping in mind that the viewers under study are children with ASD and are most likely high systemizers, Baron-Cohen's work tells us that if content creators can make the learning of emotions systematic in some way, these children have a greater chance of learning the material. So if the content contained in shows like *Thomas* is germane to the emotion-recognition lesson, makes it easy for viewers to link what they are learning to material they already know, keeps a small narrative distance in the storyline, and presents relevant material in both words and pictures, this "hybrid theory" predicts that the program can successfully utilize the capacity of working memory to teach emotion recognition to children with ASD. Obvious next steps then include detailed content analyses of episodes of *Thomas* and similar programs, specifically looking for the elements of germane cognitive load, opportunities for links, narrative distance, and dual-channel instruction proposed in this new theory, again working under the assumption that this special population of child viewers are most likely high systemizers. Assuming all of these elements are present in the content, experimental pretest-posttest emotion-recognition research is the ultimate goal.

10. References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, fifth edition.* Arlington, VA: American Psychiatric Publishing.

Baddeley, A.D. (1986) *Working memory*. Oxford: Clarendon Press.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Baron-Cohen, S. Golan, O., Chapman, E., Granader, Y. (2007). Transported into a world of emotion. *The Psychologist, 20*(2), 76-77.

Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R., Ashwin, C., Billington, J.et al. (2005). Empathizing and systemizing in autism spectrum conditions. In Volkmar, F., Klin, A, & Paul, R. (Eds.), *Handbook of Autism and Pervasive Developmental Disorders*. New York: John Wiley and Sons, Inc.

Baron-Cohen, S. (2008). Autism, hypersystemizing, and truth. *The Quarterly Journal of Experimental Psychology, 61* (1), 64-75.

Basil, M. D. (1994). Multiple resource theory I: Application to television viewing. *Communication Research, 21*(2), 177-201

de Koning, B.B., Huib, K.T., Remy M.J.P. Rikers, & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*, 731-746.

Charlop, M. H., & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied Behavior Analysis, 22,* 275-285.

Charlop-Christy, M.H. & Daneshvar, S. (2002). Using video modeling to teach perspective taking to children with autism. *Journal of Positive Behavior Interventions, 5*(1), 12-21.

Charlop, M. H., & Walsh, M. E. (1986). Increasing autistic children's spontaneous verbalizations of affection: An assessment of time delay and peer modeling procedures. *Journal of Applied Behavior Analysis, 19,* 307-314.

Corbett, B.A. (2003). Video modeling: a window into the world of autism. *The Behavior Analyst Today, 4,* 367-75.

Christopher, J. A., Sears, L. L., Williams, P. G., Oliver, J., & Hersh, J. (2004). Familial, medical, and developmental patterns of children with autism and a history of language regression. *Journal of Developmental & Physical Disabilities, 16*(2), 163-170.

Collins, W. A., Sobol, B. L., & Westby, S. (1981). Effects of adult commentary on children's comprehension and inference from a televised dramatic narrative. *Child Development, 49*, 389-399.

Crawley, A. M., Anderson, D. R., Santomero, A., Wilder, A., Williams, M., Evans, M. K., et al. (2002). Do children learn how to watch television? The impact of extensive experience with Blue's Clues on Preschool Children's Television Viewing Behavior. *Journal of Communication, 52*, 264-280.

D'Ateno, P., Mangiapanello, K., Taylor, B.A. (2003). Using video modeling to teach complex play sequences to a preschooler with autism. *Journal of Positive Behavior Interventions,5,* 5-11.

Dawson, G., Toth, K., Abbott, R., Osterlig, J., Muson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*, 271-183.

Dickstein, S., & Parke, R. D. (1988). Social referencing in infancy: A glance at fathers and marriage. *Child Development, 59*, 506-511.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63-100.

Dorwick, P.W., & Jesdale, (1991). *Practical Guide to Using Video in the Behavioral Sciences*. John Wiley & Sons, Inc. New York. p. 75.

Durand, V. M. (2005). Past, present, and emerging directions in education. In D. Zager (Ed.), *Autism spectrum disorders: Identification, education, and treatment* (3rd ed., pp. 89-109). Mahwah, NJ: Lawrence Erlbaum.

Eysenck, M. (1993). *Principles of cognitive psychology*. Hillsdale, NJ: Erlbaum.

Fisch, S. M. (2004). *Children's learning from educational television: Sesame Street and beyond*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fisch, S. M. (2009). Educational television and interactive media for children. In J. Bryant & M.B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 402-435). NewYork: Routledge.

Gabriels, R. L., Cuccaro, M. L., Hill, D. E., Ivers, B.J., & Goldson, E. (2005). Repetitive behaviors in autism: Relationships with associated clinical features. Research in *Developmental Disabilities, 26,* 169-181.

Gathercole, S. E., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman & J. E. Thomas (Eds.), *Handbook of perception and human performance* (pp. 41-48). New York: Wiley.

Haney, M. R. (2013). *Understanding children with Autism Spectrum Disorders: Educators partnering with families*. Los Angeles: Sage.

Hart, S. G. (1986). Theory and measurement of human workload. In J. Zeidner (Ed.), *Human productivity enhancement* (pp. 396-455). New York: Praeger.

Heflin, L. J., & Aliamo, D. F. (2007). *Students with Autism Spectrum Disorder*. Upper Saddle River, NJ: Pearson.

Hoffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning And Instruction, 17*(6), 722-738.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child, 2*, 217-250.

Kinney, E.M., Vedora, J. & Stromer, R. (2003). Computer-presented video models to teach generative spelling to a child with an autism spectrum disorder. *Journal of Positive Behavior Interventions, 5,* 22-29.

Kwaitek, K., & Watkins, B. (1981). *The systematic viewer: An inquiry into the grammar of television* (First annual report to the Spenser Foundation). Ann Arbor, MI: Children's Media Project, University of Michigan

Lang, A. (1992). A limited capacity theory of television viewing. Paper presented at the Information Systems Division of the International Communication Association in Miami, FL.

Lang, A. (2000). The limited capacity model of mediate message processing. *Journal of Communication, 50(*1), 46:70.

Lang, A. (2006). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. *Journal of Communication, 56*, S57: S80.

Lang, A., & Basil, M. D. (Eds.). (1998). Attention, resource allocation and communication research: What do secondary task reaction times measure anyway? (Vol. 21). Beverly Hills, CA.: Sage.

Lawson, M. J., & Kirby, J. R. (1981). Training in information processing algorithms. *British Journal of Educational Psychology, 51*, 321-355.

Leahy, W. & Sweller, J. (Ed.). (2011). *The Psychology of Learning and Motivation* Volume 55 (chapter 2). San Diego, CA: Elsevier Inc.

LeBlanc, L.A., Coates, A.M., Daneshvar, S., Charlop-Christy, M.H., Morris, C. & Lancaster, B.M. (2003). Using video modeling and reinforcement to teach perspective-taking skills to children with autism. Journal of Applied Behavior Analysis. 36, 253-57.

Mayer, R. E. (2001). *Multi-media learning*. Cambridge, UK: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: Cambridge University Press.

Mayer, R.E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*, 312-320.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive Load Theory: Instructional implications of the interation between information structures and cognitive architecture. *Instructional Science, 32*, 1-8.

Salomon, G., & Leigh, T. (1984). Predispositions about learning from television and print. *Journal of Communication, 34*, 119-135.

Shipley-Benamou, R., Lutzker, J.R., & Taubman, M. (2002). Teaching daily living skills to children with autism through instructional video modeling. *Journal of Positive Behavior Interventions, 4,* 165-176.

Shriffin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and general theory. *Psychological Review, 84*, 127-190

Sorden, S.R. (2012). The cognitive theory of multimedia learning. In B.J. Irby, G. Brown & R. Lara-Alecio (eds). *Handbook of educational theories*. Charlotte, NC: Information Age Publishing.

Stirling, A. (2007). *Making connections: A report on the special relationship between children with autism and Thomas & Friends*. London: The National Autistic Society.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12 (2): 257–285.

Taylor, B.A., Levin, L., & Jasper, S. (1999). Increasing play-related statements in children with autism toward their siblings: effects of video modeling. *Journal of Developmental and Physical Disabilities, 11*, 253-264.

Wert, B.Y. & Neisworth. J.T. (2003). Effects of video self-modeling on spontaneous requesting in children with autism. *Journal of Positive Behavior Interventions, 5,* 30-36