2017

Comprehension Models of Audiovisual Discourse Processing

Courtney Anderegg
George Fox University, canderegg@georgefox.edu

Fashina Alade
Northwestern University

David R. Ewoldsen
Michigan State University

Zheng Wang
The Ohio State University

Comprehension Models of Audiovisual Discourse Processing

Courtney Anderegg[1], Fashina Aladé[2], David R. Ewoldsen[3], & Zheng Wang[1]

1 School of Communication, The Ohio State University, Columbus, OH 43210, USA
2 Department of Communication Studies, Northwestern University, Evanston, IL 60208, USA
3 Department of Media & Information, Michigan State University, East Lansing, MI 48824, USA

Comprehension is integral to enjoyment of media narratives, yet our understanding of how viewers create the situation models that underlie comprehension is limited. This study utilizes two models of comprehension that had previously been tested with factual texts/videos to predict viewers' recall of entertainment media. Across five television/film clips, the landscape model explained at least 29% of the variance in recall. A dual coding version that assumed separate verbal and visual representations of the story significantly improved the model fit in four of the clips, accounting for an additional 15–29% of the variance. The dimensions of the event-indexing model (time, space, protagonist, causality, and intentionality) significantly moderated the relationship between the dual coding model and participant recall in all clips.

Keywords: Comprehension, Landscape Model, Event-Indexing Model, Entertainment Media, Situation Models, Cognitive Psychology.

doi:10.1111/hcre.12107

When people watch television or read the newspaper, an implicit goal is to comprehend what they are watching or reading. Comprehension involves integrating existing knowledge with new information to form a coherent mental representation of interconnected concepts and ideas. This mental representation is called a situation model (Wyer, 2004). There is a fairly large body of research on written texts that demonstrates that people construct situation models during reading (e.g., Lee, Roskos-Ewoldsen, & Roskos-Ewoldsen, 2008). However, with few exceptions, the study of how people create a coherent understanding of audiovisual media has been overlooked (Busselle & Bilandzic, 2008; Magliano, Miller, & Zwaan, 2001; Roskos-Ewoldsen, Roskos-Ewoldsen, Yang, & Lee, 2007). This is problematic because comprehension processes play a critical role in how media texts are processed (Busselle & Bilandzic, 2008) and in the consequences of media consumption (Roskos-Ewoldsen, Davies, & Roskos-Ewoldsen, 2004; Yang & Roskos-Ewoldsen, 2007). With television and

Corresponding author: Courtney Anderegg; e-mail: anderegg.4@osu.edu

movies occupying the majority of Americans' time spent with entertainment media (Nielsen, 2015), it is important to extend the study of comprehension and situation models to these mediums.

The primary goal of this research is to determine if the landscape model (van den Broek, Risden, & Husebye-Hartman, 1995), originally conceptualized for written text, can be applied to audiovisual clips. Second, dual coding theory (DCT; Paivio, 1986, 1991), which assumes separate verbal and visual representations of a narrative, will be examined as an extension of the landscape model (Lee et al., 2008). Finally, the event-indexing model (Zwaan, Langston, & Graesser, 1995) will be explored as a way to augment the landscape model's predictive power.

Understanding how people comprehend stories told by the media, with a focus on mental representations, lends important insight into many different media outcomes (Busselle & Bilandzic, 2008; Roskos-Ewoldsen et al., 2004; Yang & Roskos-Ewoldsen, 2007). For example, if comprehension models are able to predict what aspects of a story will be understood and stored in long-term memory for retrieval, this information may be used to create more effective entertainment programming that aims to educate audiences on certain issues (e.g., entertainment-education; Moyer-Gusé, 2008) or persuade audiences to consider certain viewpoints or behaviors (e.g., Slater, 2002). In addition, narratives are often complex (e.g., multiple characters, several storylines, shifts in time, or place) in an attempt to achieve goals such as audience engagement and enjoyment. Applying comprehension models to audiovisual narratives could increase our understanding of how complex information is represented in memory and may shed light on what aspects of the story structure resonate most with viewers (Busselle & Bilandzic, 2008).

The landscape model of comprehension

One of the most successful models of text comprehension is the landscape model (van den Broek et al., 1995), which was developed to explain how people generate a coherent understanding of a story and represent the story in memory. In particular, the model focuses on how concepts within a story are activated in memory during reading. The landscape model uniquely focuses on coherence by looking at the relationship between the online processing of a story and the memorial representation of that story. By examining participant memory for a text, the landscape model takes advantage of the well-established finding within cognitive psychology that a greater level of activation of a particular concept during reading results in better memory for that concept (Lee et al., 2008; van den Broek, Risden, Fletcher, & Thurlow, 1996; van den Broek, Young, Tzeng, & Linderholm, 1999). Thus, by examining an individual's memory for concepts embedded in a text, one can test the theory's predictions concerning the mental representation that results from concept activation during the processing of the story.

The landscape model argues that there are several sources of activation that occur during the comprehension process (van den Broek et al., 1996, 1999). First, concepts within the current sentence in a book or scene in a movie will be activated. Second, because activation dissipates across time (Higgins, Bargh, & Lombardi, 1985), central concepts from the immediately preceding sentence or scene should still be activated, albeit at a lower level of activation. Third, concepts from earlier in the story may be reactivated when they are necessary for maintaining story coherence. Fourth, world knowledge that is necessary for understanding the story will be activated. According to the landscape model, the cognitive representation of the story will reflect these four sources of activation.

The landscape model has been shown to predict participant memory for text-based stories quite well (Lee et al., 2008; Tzeng, 2007; van den Broek & Gustafson, 1990). Three attributes of the activated concepts are used by the landscape model to predict memory for a story: the number of times a concept is activated, the degree to which a concept is activated across all meaningful units of a story (e.g., scenes or sentences), and the associations formed between simultaneously activated concepts. According to the model, each of these attributes reflects a unique aspect of the mental representation of a story. First, the number of times a concept is activated within a story is expected to influence recall, because a memory trace of a concept will be created each time a concept is activated. A greater number of memory traces generated for a concept leads to a greater likelihood that the concept will be recalled later. Second, the degree to which a concept is activated across all meaningful units of the story is expected to influence recall of that concept, because concepts with higher levels of overall activation have stronger memory traces created and stored in memory. Finally, associations between concepts that are activated simultaneously should influence the likelihood that a concept is recalled (van den Broek et al., 1995). This is because two concepts that are activated simultaneously will be linked in memory; it is assumed that concepts with more linkages in memory are more likely to be recalled than those concepts with fewer linkages. While the model assumes that these three attributes of activation can theoretically operate independently of each other, pragmatically, they tend to highly correlate ($r > .80$), and the independent contribution of each is difficult to ascertain. However, the overall predictions of the landscape model can be tested (Lee et al., 2008).

The landscape model started as a comprehension model for text and has been successfully applied to reading of entertainment (e.g., Linderholm & van den Broek, 2002) and scientific texts (e.g., van den Broek, Kendeou, Sung, & Chen, 2003). More recently, it has been extended to the comprehension of audiovisual discourse, such as to predict memory of brand placement in films (Yang & Roskos-Ewoldsen, 2007) and recall of short television news stories (Lee et al., 2008). However, it has not yet been applied to entertainment audiovisual media.

One weakness of the landscape model is that it does not address the role of inferences, making it unclear whether the model can predict viewers' memory for entertainment stories (Lee et al., 2008). Research suggests that entertainment stories rely on more inferences for comprehension than factual stories (Graesser, Singer, & Trabasso, 1994; Lee, Roskos, & Ewoldsen, 2013), and that participants make several types of inferences while reading (Trabasso & Magliano, 1996) and watching movies (Lee et al., 2013). If the landscape model is able to predict memory for entertainment narratives, this would suggest that inferences, while important for comprehension, may not play a central role in mental representations of those stories. Because prior research has successfully applied the model to reading for entertainment (e.g., Linderholm & van den Broek, 2002) and to audiovisual media (e.g., Lee et al., 2008), we expect:

H1: The landscape model will significantly predict memory for audiovisual entertainment narratives.

Another primary goal of this study is to extend the landscape model by incorporating DCT (Paivio, 1986, 1991). DCT predicts that visual and auditory information have independent yet interacting representations in memory (Paivio, 1986, 1991). The landscape model was developed as a model for text comprehension, assuming a single representation of a story. However, comprehension in situations other than reading often relies on both visual and auditory information. Although the landscape model would predict comprehension of an audiovisual narrative in accordance to concept activation, it does not delineate between concepts conveyed as audio versus video representations. In applying the landscape model to television news stories, Lee et al. (2008) demonstrated that the model had to assume both a verbal and a visual representation. The dual coding landscape model accounted for 73% of the variance in participant recall of news information, whereas the original model accounted for only 32% of the variance (Lee et al., 2008). To date, this is the only test of the dual coding model.

There is reason to believe that the dual coding model should also work well with audiovisual entertainment narratives. In these narratives, one channel (i.e., visual or verbal) is usually dominant in conveying information to the viewer. For example, a character may use nonverbal (visual only) communication to indicate emotion, or we may hear a character speak (audio only) and add value to the plot. In these instances, the dual coding landscape model would capture the independent contributions of the auditory and visual elements of the entertainment narrative, which may increase memory of certain story concepts. Therefore, we predict:

H2: The dual coding landscape model, in comparison to the original landscape model, will better predict memory for audiovisual entertainment narratives.

The event-indexing model

In addition, this study begins to explore the integration of the landscape model and the event-indexing model (Zwaan, Langston, et al., 1995). Although the landscape model examines levels of concept activation within a story, it is silent as to the types of concepts likely to be represented within a situation model. The event-indexing model provides an important complement to the landscape model by proposing five critical dimensions of information that play a role in comprehension: time, space, protagonist, causality, and intentionality (Magliano et al., 2001; Zwaan, Langston, et al., 1995; Zwaan, Radvansky, Hilliard, & Curiel, 1998).

Comprehension via the event-indexing model relies on the reader creating a situation model of a narrative and updating the model based on shifts in multiple dimensions. Once a situation model is created in memory, shifts within the narrative along one or more of the five event-indexing dimensions will prompt the reader to update the situation model in working memory. Updating a situation model based on changes along the five dimensions has been found to aid readers in comprehending a text (Zwaan, Magliano, & Graesser, 1995; Zwaan et al., 1998). Thus, the event-indexing model has often been examined in terms of event boundaries. The boundary that occurs between two units represents a portion of the narrative where some or all of the event-indexing dimensions shift. Prior empirical studies have used segmentation tasks to divide the narrative and analyze discontinuities in the five dimensions (e.g., Cutting, 2014; Magliano et al., 2001). However, because this study uses the event-indexing model to refine the landscape model, we do not take a segmentation approach to the event-indexing analysis. Instead, we examine how identifying concepts relevant to the five dimensions of the event-indexing model can extend the landscape model's ability to predict

As noted previously, many empirical tests of the event-indexing model have focused on text processing (Zwaan & Radvansky, 1998). However, researchers have begun to extend this model to other modes of story processing, including film viewing. For example, Magliano et al. (2001) found that film viewers were sensitive to shifts of time, place, and/or characters' actions. In another study, participants were asked to rate the similarity between events within a story. Results showed that the five dimensions of the event-indexing model influenced participant judgments (Zwaan, Magliano, et al., 1995). These studies provide evidence as to the importance of the event-indexing model dimensions in comprehension.

This study begins to explore a model that combines the landscape model and the event-indexing model to predict participant memory for entertainment clips. Zwaan (1999) proposed the *dimensional equality assumption*, which suggests that each of the five dimensions of the event-indexing model are equally important in comprehension processes. However, in their empirical studies, Zwaan and colleagues found that the five dimensions may not always play an equally important role in comprehension (Magliano et al., 2001; Zwaan, Magliano, et al., 1995; Zwaan & Radvansky, 1998; Zwaan et al., 1998). These studies demonstrated that one or more of the dimensions may play a particularly important role in conjunction with a reader's goals. For example, motivation may be important within a murder mystery because the motivation for catching the killer and/or the motivation for murdering someone may drive processing of the story. In this case, motivation should play a more important role in predicting memory for the story than the other dimensions. In other words, in certain stories or genres, one or more of the dimensions may disproportionately influence an individual's memory for the story. In terms of merging the event-indexing and landscape models, we are interested in understanding whether the event-indexing dimensions provide insight into the situation models that viewers construct during comprehension beyond that explained by the landscape model. Thus, we ask the following:

RQ1: Will categorizing concepts into the event-indexing dimensions significantly moderate the relationship between the landscape model and participant recall?

Method

Participants and procedures

The researchers collected data for this study at two different points in time. For the first phase of data collection, each participant ($N = 74$) watched three television/film clips (i.e., Clip A — *That '70s Show*, Clip B — *Matlock*, and Clip C — *Once Upon a Time in the West*); the order was randomized between participants. At the start of the experiment, researchers instructed the participant to watch the clips as he or she would normally watch television. Following the viewing of all clips, cued recall was measured. The participant was presented with a frame of the initial shot of the clip as the cue and asked to write down as much as he or she could remember from the clip.

For the second phase of data collection, two additional television clips were examined in order to extend the generalizability of the findings to more current television programming. Each participant ($N = 245$) was randomly assigned to watch one of two television clips (n [Clip D — *Bones*] = 120; n [Clip E — *Parenthood*] = 125). At the start of the experiment, we instructed the participant to watch the clip as he or she would normally watch television. Following viewing, cued recall was measured. Both phases of data collection occurred at a large university in the Midwestern United States where undergraduate students were awarded course credit for participation.

Stimuli

In total, we selected five audiovisual clips for use in this study. The clips were selected from a variety of genres and time periods to ensure that results would be generalizable. Clip A was a 2 minute 38 second clip from *That '70s Show* (Season 1, Episode 20: "A New Hope"), a sitcom that aired from 1998 to 2006. In the clip, the male protagonist confronts his girlfriend about time she spent with another man, and then consults his friends about the argument. Clip B was a 2 minute 44 second clip from the legal drama *Matlock* (Season 3, Episode 6: "The Captain"), which aired from 1986 to 1996. In this clip, a detective confronts the captain of a police force about his involvement in a series of murders. Clip C was a 2 minute 42 second clip from the film *Once Upon a Time in the West*, a 1968 western. In this clip, a widow confronts a laundryman about a plot to take her late husband's land away, and the laundryman relays the conversation to his superiors while followed by a stranger. Clip D was a 1 minute 26 second clip from *Bones* (Season 7, Episode 9: "The Don't in the Do"), a procedural drama that premiered in 2005 and still airs currently. In this clip, a forensic mystery-solving couple is getting ready for work when the woman reveals she is uncomfortable in her post-baby body, and her partner fails to comfort her. Clip E was a 56-second clip from *Parenthood* (Season 3, Episode 5: "Nora"), a family drama that aired from 2010 to 2015. In this clip, the female protagonist brings home a guest and invites her to spend the night, which is a surprise to her husband.

Variables and coding

Initial studies examining the landscape model with print narratives used a sentence as the unit of analysis (van den Broek et al., 1995). In audiovisual narratives, the event-indexing model examines dimensional shifts and boundaries using a shot as the unit of analysis (Magliano et al., 2001). Because one of the main goals of the study was to explore how the event-indexing model can augment the landscape model's predictions, we needed to define a unit of analysis that could reasonably map on to each model's specifications. In addition, because research suggests that the total time an item is in working memory predicts recall from long-term memory, we felt it important for each segment to be equal in length. Therefore, the unit of analysis for this study was a 4-second segment of the clip.

Before coding began, the researchers identified the important concepts in each clip. A "concept" was defined as a word or phrase that represented information essential to understanding (Lee et al., 2008). The five dimensions of the event-indexing model (i.e., time, space, protagonist, intentionality, and causality) were used as a guide in deriving concepts. The researchers watched each clip together and identified concepts representing protagonists (e.g., Booth), space (e.g., locker room), time (e.g., night), intentionality (e.g., flirt), and causality (e.g., find medal). Forty-eight concepts were identified for Clip A, 44 for Clip B, 50 for Clip C, 44 for Clip D, and 46 for Clip E.

Then, the coders carried out three separate coding processes. First, they coded concept activation for each clip based on procedures for testing the landscape model (e.g., Lee et al., 2008; van den Broek et al., 1995). This process was used to code concepts according to the original landscape model and the dual coding landscape model. Next, the coders categorized each concept in a clip into one of the five dimensions identified by the event-indexing model (i.e., time, space, protagonist, causality, and intentionality; Zwaan, 1999). Finally, they coded participant recall to identify the concepts recalled when participants described the narrative from memory.

Concept activation

Van den Broek et al. (1995) and Lee et al. (2008) outline the steps for coding concept activation for print narratives according to the landscape model. Due to the inclusion of thousands of possible concepts in audiovisual entertainment media (e.g., background characters, scenery, objects in the environment), it was necessary to reframe the coding steps to more specifically account for concept activation in both the original and the dual coding version of the landscape model. In the original landscape model coding scheme (van den Broek et al., 1995), each concept is assigned one activation score per unit of analysis. This study defined dimensions of *presence*, *importance*, and *explicitness* (explained below) based on the original coding scheme to help coders differentiate the activation levels of concepts.

Trained graduate students coded concept activation after establishing intercoder reliability (Krippendorff's α ranged from .76 to 1.00 for all coding items). Coders viewed the clip in 4-second segments and coded concept activation within each segment in terms of three dimensions of activation. *Presence* measured if and how the concept was present within the segment (0 = not present, 1 = present, background, and 2 = present, foreground or directly involved in action). *Importance* measured if a concept was important to moving the plot forward (0 = not at all important, 1 = deactivating in importance, and 2 = important). The deactivation code captures a gradual decline in activation if a concept is not reinstated in a subsequent unit of analysis (van den Broek et al., 1996). If it is not reinstated, the activation value reduces to half of the original value (in this case, from 2 to 1) in the subsequent unit of analysis and returns to zero in the third unit of analysis. *Explicitness* measured whether the concept was implicit (i.e., alluded to or inferentially referenced) or explicit (i.e., clearly stated or seen; 0 = not applicable or not present, 1 = implicit, and 2 = explicit). Thus, a concept could have an activation score from 0 to 6 in each 4-second segment of the clip. In order to test the original and dual coding landscape model, concept activation levels were coded for three versions of each clip: video only (i.e., only picture, no sound), audio only (i.e., only sound, no picture), and audiovisual.

After coding concept activation, the researchers formed three sets of activation vectors for each version of a clip (i.e., video only, audio only, and audiovisual), totaling nine activation vectors per clip. A set of activation vectors included: *degree* of activation, *number* of activations, and *association* of activation (Lee et al., 2008; van den Broek et al., 1995). *Degree* of activation refers to each concept's activation score per 4-second segment summed across the entire clip. Each concept's activation score could range from 0 to 6 per 4-second segment. *Number* of activations refers to the number of 4-second segments in the clip in which the concept was activated to any degree. Concepts activated in a segment were coded as 1 and summed across all segments. *Association* of activation refers to the strength of each concept's total activation score when viewed in terms of simultaneously activated concepts. This measure of activation is a good predictor of memory, because concepts that are simultaneously activated tend to be connected in memory (van den Broek et al., 1995; Wyer, 2004). In order to create this vector, the activation score of each concept per segment was multiplied by the activation score of each of the other concepts identified in the clip for that segment. Thus, scores for association of activation could have ranged from 0 to 36 for each concept × concept association; this number was summed across all concept × concept associations for each concept per segment and then across all segments of the clip to create an activation score for each concept across the entire clip.

Concept categorization per event-indexing model dimensions

After establishing intercoder reliability (Krippendorff's $\alpha = 1.00$ for all five clips), each concept was categorized as representing one of the event-indexing dimensions: time, space, protagonist, causality, or intentionality. Given the short length of the clips, neither time nor space had much variability when considered alone. These dimensions tend to account for the least variance in the comprehension of text (Zwaan & Radvansky, 1998) and are frequently found in combination within visual narratives (Cutting, 2014). Thus, time and space were collapsed into one category.

Table 1 Means and Standard Deviations for the Percentage of Participants that Recalled a Concept and the Average Number of Times a Concept Was Mentioned in Cued Recall per Participant

Television Show/Film	Concepts in Clip *N*	Percentage of Participants that Recalled a Concept *M*% (*SD*)	Average Number of Times a Concept Was Recalled per Participant *M* (*SD*)
That '70s Show	48	26.86 (28.13)	0.62 (1.42)
Matlock	44	24.53 (22.67)	0.55 (1.32)
Once Upon a Time in the West	50	18.89 (24.04)	0.39 (0.84)
Bones	44	16.21 (25.96)	0.35 (0.96)
Parenthood	46	17.86 (26.11)	0.35 (0.80)

Cued recall of concepts

After establishing intercoder reliability (Krippendorff's α ranged from .76 to .93), participant recall was coded (number of words, $M = 68.61$, $SD = 43.58$). For each clip, coders identified whether or not a participant mentioned each of the concepts in his or her response. For example, if a participant recalling the story events from *That '70s Show* wrote, "The guy gave Donna advice about a story and Donna was excited," coders would indicate that David (i.e., the guy), Donna, advice, story, and excited were mentioned in this response by giving those concepts a score of 1. Concepts not mentioned received a 0. Only concepts included in the concept list developed by the researchers were coded. Scores were summed across participants and divided by the total number of participants in order to create a variable quantifying the percentage of participants that recalled the concept.

In addition, coders identified the number of times that a participant mentioned each concept in his or her response. For instance, in the previous example, coders would count each concept mentioned once except for Donna, which was mentioned twice (e.g., gave Donna advice, Donna was excited). Coders summed the number of times a concept was mentioned across participants and divided it by the total number of participants to create a variable quantifying the average number of times a participant mentioned a concept in a response. Means and standard deviations for the two cued recall measures across participants are shown in Table 1.

Results

The original versus dual coding landscape model

Following the analytical strategies of prior studies (Lee et al., 2008; van den Broek et al., 1995), hierarchical regression was used to compare the original (i.e., audiovisual) and dual coding (i.e., separate audio and video) landscape models. The three activation vectors (i.e., degree, number, and association of activation) of the original landscape model were entered as the first step to predict recall as measured by the percentage of participants that recalled a concept. The dual coding landscape model was entered as the second step. This model includes three activation vectors for the audio-only and video-only versions of each clip (six vectors total).

H1 posited that the landscape model would predict participant memory for audiovisual entertainment stories. The original landscape model significantly explained between 29 and 66% of the variance in participants' recall of the concepts across the five clips, as indicated by the regression R^2 estimates (see Table 2 for model statistics). The predictors play an important role in the overall fit of the model, but individual contributions of each predictor cannot be ascertained due to the high levels of multicollinearity (Cohen, Cohen, West, & Aiken, 2003). The correlation between the three activation vectors across the five clips ranged from .74 to .99. As in previous studies testing the landscape model (Lee et al., 2008), because we are interested in the overall model fit and the combined — not individual — contribution of the predictors, multicollinearity among predictors is not an issue. Thus, H1 was supported.

H2 predicted that the dual coding version of the landscape model would outperform the original landscape model. In four of the five clips, the dual coding model produced a significant increase (all $p\text{s} < .05$) in variance explained (from 15 to 29%) for participant recall as compared to the original model (see model statistics in Table 2). The dual coding landscape model did not significantly outperform the landscape model in Clip C (*Once Upon a Time in the West*). Thus, the dual coding landscape model can predict memory for entertainment audiovisual stories and provides a significantly better fit than the original landscape model in four of the five clips, partially supporting H2.

The above analyses used the percentage of concept recall across participants as the dependent variable. However, these analyses can also be conducted using the average number of times a concept was recalled. The same pattern of significant effects was found when using the second measure, but the R^2 estimates for the variance explained in the landscape model ranged from 63 to 85%, and the dual coding landscape model continued to outperform the original landscape model in four of the five clips (increased variance explained ranged from 6 to 29%).

The event-indexing model

To explore RQ1, each of the event-indexing model dimensions (i.e., time/space, protagonist, causality, and intentionality) was examined for each clip. Due to the exploratory nature of the analyses, we examined the interaction effects of the event-indexing dimensions on the relationship between the dual coding landscape model and participant recall. Because individual contributions of activation variables cannot be ascertained due to multicollinearity, only the association variable was used in this analysis based on prior research demonstrating it is a good predictor of memory (van den Broek et al., 1995; Wyer, 2004). Because the dual coding landscape model accounted for more of the variance in recall than the original model in four

Table 2 Regression Statistics Predicting Concept Recall from the Original (Step 1) and Dual Coding (Step 2) Landscape Model

	Clip A *That '70s Show*		Clip B *Matlock*		Clip C *Once Upon a Time in the West*		Clip D *Bones*		Clip E *Parenthood*	
Predictors	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Step 1 (Audiovisual)										
Degree	0.98	0.54	−1.91*	0.81	−0.07	0.44	2.73*	1.12	−0.68	2.27
Number	−5.32	2.68	0.67	0.74	−0.52	1.37	−1.48	1.76	−4.79	4.10
Association	0.02*	0.01	−0.00	0.02	0.02	0.01	−0.04	0.03	0.04	0.03
Model fit	$R^2 = .48$, $F(3, 44) = 13.60$, $p < .0001$		$R^2 = .29$, $F(3, 40) = 5.48$, $p = .003$		$R^2 = .52$, $F(3, 46) = 16.63$, $p < .0001$		$R^2 = .66$, $F(3, 40) = 26.19$, $p < .0001$		$R^2 = .44$, $F(3, 42) = 11.00$, $p < .0001$	
Step 2 (Audio only)										
Degree	6.22	4.47	−0.13	9.31	0.56	0.98	−2.92	7.48	−17.14**	5.83
Number	−0.65	1.24	0.73	1.65	−0.84	3.90	0.03	1.69	7.17*	3.95
Association	−0.01	0.03	−0.02	0.04	0.01	0.03	0.05	0.04	0.06	0.06
Step 2 (Video only)										
Degree	2.19	4.99	−5.16	4.04	−0.04	1.06	2.31	5.96	3.52	6.90
Number	−0.47	1.24	1.57	1.27	7.29*	3.29	−2.54	1.63	2.58	2.12
Association	0.03	0.02	−0.01	0.03	−0.04	0.03	0.05	0.08	−0.06	0.03
Model fit	$\Delta R^2 = .18$, $F(6, 38) = 3.44$, $p = .008$		$\Delta R^2 = .26$, $F(6, 34) = 3.20$, $p = .013$		$\Delta R^2 = .08$, $F(6, 40) = 1.42$, $p = .232$		$\Delta R^2 = .15$, $F(6, 34) = 4.73$, $p = .001$		$\Delta R^2 = .29$, $F(6, 36) = 6.21$, $p < .001$	

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

of the five clips, we used the dual coding model to examine if the event-indexing dimensions would add to the variance explained in predicting recall.

In order to test for the interaction effects of the event-indexing model dimensions, it was necessary to construct a conditional effects model with a multicategorical variable (i.e., the event-indexing dimensions) as a moderator. The event-indexing dimensions were dummy coded and the time/space dimension was used as the reference group due to the limited variability of this dimension in each clip. Therefore, reported interaction effects refer to the moderating effect of the event-indexing dimension on the relationship between concept activation and participant recall when compared to the reference group of time/space. Positive interaction coefficients indicate that the dimension did a better job of predicting participants' memory relative to the time/space dimension.

The model for the audio only version of Clip A (*That '70s Show*) accounted for a significant portion of the variance in participant recall, $r = .97$, $R^2 = .94$, $F(4, 43) = 166.89$, $p < .001$. There was a significant conditional effect of concept activation on recall for concepts coded as time/space. Additionally, the interaction effect for concepts coded as the intentionality dimension was significant, meaning that concepts coded as intentionality moderated the relationship between concept activation and recall when compared to the reference group. The interaction effect for concepts coded as causality was also significant. The model for the video-only version of Clip A also accounted for a significant portion of the variance in participant recall, $r = .97$, $R^2 = .95$, $F(4, 43) = 197.50$, $p < .001$. There was a significant conditional effect of concept activation on recall for concepts coded as time/space, as well as a significant interaction effect for concepts coded as protagonist.

In the audio-only version of Clip B (*Matlock*), the model accounted for a significant portion of the variance in participant recall, $r = .74$, $R^2 = .55$, $F(3, 40) = 12.08$, $p < .001$. Although the conditional effect was not significant, $p = .074$, there was a significant interaction effect for concepts coded as causality. In the video-only version, the model accounted for a significant portion of the variance in participant recall, $r = .58$, $R^2 = .33$, $F(4, 39) = 4.88$, $p = .005$. Although the conditional effect was not significant, $p = .920$, there was a significant interaction effect for concepts coded as protagonist.

In Clip C (*Once Upon a Time in the West*), the model for the audio-only version of the clip accounted for a significant portion of the variance explained for participant recall, $r = .90$, $R^2 = .82$, $F(4, 45) = 49.97$, $p < .001$. The model showed a significant conditional effect of concept activation on recall for concepts coded as time/space. In addition, there was a significant interaction effect for concepts coded as protagonist, causality, and intentionality. In the video version of this clip, the model accounted for a significant portion of the variance explained for participant recall, $r = .88$, $R^2 = .77$, $F(4, 45) = 37.06$, $p < .001$. There was a significant conditional effect of concept activation on recall for concepts coded as time/space. In addition, there was a significant interaction effect for concepts coded as causality.

Although the model for the audio version of Clip D (*Bones*) accounted for a significant portion of the variance explained for participant recall, $r = .89$, $R^2 = .79$, $F(4, 39) = 36.65$, $p < .001$, it did not show a significant conditional effect for concepts coded as time/space, $p = .762$, nor did it show significant interaction effects for the event-indexing dimensions of protagonist, $p = .98$, causality, $p = .88$, or intentionality, $p = .91$. The model for the video version of the clip significantly explained a portion of the variance in participant recall, $r = .85$, $R^2 = .73$, $F(4, 39) = 25.96$, $p < .001$, but did not show a significant conditional effect of the time/space dimension, $p = .805$. However, the interaction effect for concepts coded as protagonist was significant.

Finally, although the models for the audio, $r = .79$, $R^2 = .62$, $F(4, 41) = 17.16$, $p < .001$, and video versions, $r = .78$, $R^2 = .61$, $F(4, 41) = 16.22$, $p < .001$, of Clip E (*Parenthood*) significantly explained a portion of the variance in participant recall, the models did not show significant conditional effects of the time/space dimension. The only significant interaction effect was in the video version for protagonists. See Table 3 for model statistics.

Discussion

How people comprehend stories is an important question for media scholars (Roskos-Ewoldsen & Roskos-Ewoldsen, 2008; Roskos-Ewoldsen et al., 2007). Comprehension is a fundamental cognitive process that is integral to a variety of related phenomena. For example, comprehension can aid in our understanding of media phenomena, such as cultivation effects (Roskos-Ewoldsen et al., 2004), effects of brand placements (Yang & Roskos-Ewoldsen, 2007), and transportation and perceptions of realism (Busselle & Bilandzic, 2008). However, little research has been conducted on the comprehension processes of users of audiovisual media. Although the landscape model has successfully captured online processing of text and predicted recall for text-based factual stories (van den Broek et al., 1996, 1999), only one study found support for the applicability of the landscape model to video stories (Lee et al., 2008). Lee et al.'s study also found that adding the assumption of Paivio's (1986, 1991) DCT substantially improved the ability of the landscape model to predict memory.

The goal of this study was to replicate and extend this earlier research. First, we tested the landscape model's applicability to entertainment media. The original model significantly predicted memory for all five clips, accounting for a minimum of 29% of the variance in recall. Second, this study offered further evidence for the dual coding representation proposed by Lee et al. (2008). While the original model was successful, the dual coding landscape model provided a significantly better fit to the recall data in four of the five clips, accounting for a minimum of 55% of the variance in participants' recall of the clip. It is unclear why one of the clips (*Once Upon a Time in the West*) did not benefit from the addition of the dual coding landscape model. Nevertheless, this finding augments the importance of attending to complexities in how information is represented in memory.

Table 3 Conditional and Interaction Effects of the Event-Indexing Model Dimensions

	Clip A *That '70s Show*			Clip B *Matlock*			Clip C *Once Upon a Time in the West*			Clip D *Bones*			Clip E *Parenthood*		
Predictors	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
Audio only															
Time/Space	0.07	0.03	2.70**	0.03	0.02	1.84	0.14	0.01	14.00***	0.01	0.33	0.31	−0.02	0.03	−0.75
Protagonist	0.03	0.03	1.24	−0.01	0.02	−0.15	−0.12	0.05	−2.62**	0.01	0.33	0.03	0.05	0.03	1.68
Intentionality	−0.07	0.03	−2.56**	−0.02	0.02	−0.23	−0.11	0.04	−2.88**	−0.04	0.32	−0.11	0.02	0.03	0.73
Causality	−0.06	0.03	−2.73*	−0.05	0.02	−2.13*	−0.84	0.03	−3.11**	−0.05	0.32	−0.15	0.02	0.03	0.64
Video only															
Time/Space	0.02	0.01	2.13*	0.00	0.00	−0.10	0.07	0.01	12.14***	−0.01	0.03	−0.25	−0.01	0.01	−0.76
Protagonist	0.13	0.01	13.47***	0.01	0.00	2.52*	−0.06	0.03	−1.95	0.15	0.03	4.46***	0.03	0.01	4.43***
Intentionality	0.01	0.08	0.12	−0.00	0.01	−0.10	−0.04	0.03	−0.09	−0.05	0.46	−0.01	−0.00	0.01	−0.08
Causality	0.01	0.01	0.69	0.00	0.02	−0.00	−0.05	0.01	−4.33***	0.03	0.04	0.09	−0.00	0.03	−0.01

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Finally, this study began exploring the potential for the event-indexing model (Zwaan, Langston, et al., 1995) to augment the predictive power of the landscape model (van den Broek et al., 1995). Although the landscape model provides precise predictions regarding the activation levels of concepts and their influence on a mental representation, the model does not address what types of concepts are likely to be represented within a situation model. The event-indexing model specifies five dimensions that are integral components of situation models: time, space, protagonist, causality, and intentionality. The time and space dimensions were combined in this study because the use of short clips did not allow for much variance along these two dimensions.

The results from this study suggest that merging the event-indexing model and the dual coding landscape model leads to a better understanding of comprehension processes than the dual coding landscape model alone. Although we found the event-indexing model dimensions to be a significant moderator in all five clips, it seems that the role of the dimensions is complex. Zwaan, Langston, and Graesser (1995) originally proposed that the five dimensions were equally important to the comprehension process, but subsequent research on the event-indexing model argued that different dimensions should be more important across different genres and should be sensitive to idiosyncrasies of a story (Magliano et al., 2001). Indeed, in this study, we see that different dimensions are more important to the comprehension process in some stories than in others. The results of the moderation analyses indicate a dynamic relationship between the dimensions and the activation levels of concepts. It may be that the importance of certain dimensions is dependent on the genre or it may be that different dimensions emerge as important at different time points throughout the story. For example, the protagonist dimension should be more important early in a story as viewers learn about the characters involved in the narrative.

Consider the clip from *That '70s Show* (Clip A), which involved a male protagonist (Eric) confronting his girlfriend (Donna) about time she spent with another man (David). The finding that the protagonist is the central component of the situation model for this story is consistent with a surface-level reading of the script. Eric witnesses only the tail end of an event where Donna accidentally spilled water on David's lap and was helping to clean it up. The dual coding landscape model does not account for the importance of the protagonists in this clip beyond that of the activation level. The addition of the event-indexing dimensions as moderators suggests that the concepts representing the protagonists are central to the representation of this narrative within a situation model in memory. This finding suggests that the combined use of the models allows us to predict which types of concepts will be central to the situation model of a story.

The ability to predict what viewers will remember from entertainment media could have important implications for media consumers, producers, and researchers alike. Comprehension is a fundamental feature of media enjoyment (Busselle & Bilandzic, 2008); thus, producers of entertainment media should have a vested interest in facilitating viewers' comprehension of their media products. Moreover, an understanding of how to make certain characters, goals, and settings more salient in the minds of the viewers (i.e., by increasing the degree of activation and coactivation of those concepts) would facilitate the production of more effective programming.

As just one example, consider the implication of these models of comprehension for entertainment-education campaigns. Entertainment-education is the strategy of incorporating educational messages into popular entertainment media with the purpose of teaching a positive message and/or encouraging viewers to perform healthy behaviors (Slater, 2002). Research has shown that entertainment-education can be more effective at persuading viewers than overtly persuasive educational messages (Moyer-Gusé, 2008). Predicting recall could be particularly useful for producers of entertainment-education programming where the primary goal is not only entertainment, but also *learning*. In these cases, recall and transfer of knowledge to new, similar contexts are prerequisites for programs to be effective in attaining their goals. Understanding how concepts are incorporated into situation models, and which concepts are most likely to be recalled, would allow writers and producers to create entertainment-education programs that leave viewers with a lasting memory of the intended message.

It is worth noting several limitations of this study. First, the video clips were relatively short. This is a common method of testing the landscape model due to the complexity of identifying and coding the concepts that form the situation model. However, future research should test the landscape model using longer stimuli to more accurately represent real-life viewing scenarios. Second, although the researchers attempted to identify all important concepts within each clip, the concept lists may not have been completely exhaustive. Future studies should pretest the clips with participants to diminish the possibility of leaving important concepts unidentified. Third, when asking a participant to recall story events that took place in a narrative, it is difficult to ascertain whether a participant recalled everything that he or she could remember or simply provided the main points. If the latter, these models have significantly predicted participant recall of the main story elements but might perform even better if participants offered more detail in their open-ended responses. Fourth, as argued in the introduction, each of the different predictors within the landscape model suggest slightly different strategies for designing messages to influence comprehension and memory for the text. The high levels of multicollinearity in this (and other) sample testing this model (e.g., Lee et al., 2008) suggests none of these strategies are more effective than another strategy for influencing comprehension and memory. Conversely, the current findings indicate that people involved with message design are presented with more flexibility for how to design the messages because the different strategies are highly correlated and are likely to have similar impacts on message comprehension and memory.

In addition, this study utilized a new adaptation of the landscape model coding method (Lee et al., 2008; van den Broek et al., 1995). Although the researchers felt that the adaptation used in this study was more apt to capture the complexities of television and film narratives, more research is needed to ascertain the validity of this coding scheme. The strength of this coding scheme is that it did allow for a more objective analysis of concept activation in an audiovisual platform. By improving how activation is captured within a television show or film, we are better able to capture concepts that are activated in one format or through audio and video components.

Finally, this study included an exploratory venture to determine whether concepts categorized according to the five event-indexing model dimensions could augment the landscape model's predictions. Our analysis of the event-indexing model dimensions does not conform to standard practice for understanding these phenomena (i.e., segmentation tasks/analysis). However, our intention for this study was not to view the event-indexing model in terms of event segmentation, but rather to look at the event-indexing model as a way to inform what types of concepts should be included in the landscape model. Based on the results from this exploratory analysis, our future work will continue to investigate the integration of these two models by examining the event-indexing model through a more standard segmentation task and analysis.

This study takes an important step in providing a more complete theoretical understanding of how individuals comprehend narratives, which can broaden our insight into many communicative practices. Particularly in media research, comprehension models can help tease apart the underlying mechanisms at work during viewing experiences, including, as we can now see, in the context of entertainment media. The dual-coding landscape model is shown to be a highly successful way to understand the comprehension of media texts within entertainment media by taking audio-only and video-only components of the narrative into account. In addition, although more research is needed to further develop this prospect, the innovative integration of two comprehension models (in this case the dual coding landscape model and the event-indexing model) creates the potential for even greater explanatory power and brings us closer to a more complete understanding of what goes on in the mind of the media user.

References

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, **18**, 255–280. doi:10.1111/j.1468-2885.2008.00322.x

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression: Correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cutting, J. E. (2014). Event segmentation and seven types of narrative discontinuity in popular movies. *Acta Psychologica*, **149**, 69–77. doi:10.1016/j.actpsy.2014.03.003

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, **101**, 371–395. doi:10.1037/0033-295X.101.3.371

Higgins, E. T., Bargh, J. A., & Lombardi, W. J. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 59–69. doi:10.1037/0278-7393.11.1.59

Lee, M., Roskos, B., & Ewoldsen, D. R. (2013). The impact of subtitles on comprehension of narrative film. *Media Psychology*, **16**, 412–440. doi:10.1080/15213269.2013.826119

Lee, M., Roskos-Ewoldsen, B., & Roskos-Ewoldsen, D. R. (2008). Applying the Landscape Model to comprehending discourse from TV news stories. *Discourse Processes*, **45**, 519–544. doi:10.1080/01638530802359566

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, **94**, 778–784. doi:10.1037/0022-0663.94.4.778

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology*, **15**, 533–545. doi:10.1002/acp.724

Moyer-Gusé, E. (2008). Toward a theory of entertainment persuasion: Explaining the persuasive effects of entertainment-education messages. *Communication Theory*, **18**, 407–425. doi:10.1111/j.1468-2885.2008.00328.x

Nielsen. (2015). *Home entertainment consumer trends: Digital transition tracker report*. Retrieved from http://www.nielsen.com/us/en/insights/reports/2015/home-entertainment-consumer-trends-digital-transition-tracker-report-may-2015.html

Paivio, A. (1986). *Mental representations*. New York, NY: Oxford University Press.

Paivio, A. (1991). *Images in mind: The evolution of a theory*. Hertfordshire, England: Harvester Wheatsheaf.

Roskos-Ewoldsen, B., Davies, J., & Roskos-Ewoldsen, D. R. (2004). Implications of the mental models approach for cultivation theory. *Communications*, **29**, 345–363. doi:10.1515/comm.2004.022

Roskos-Ewoldsen, D. R., & Roskos-Ewoldsen, B. (2008). Message processing. In C. R. Berger, M. E. Roloff & D. R. Roskos-Ewoldsen (Eds.), *The handbook of communication science* (pp. 129–144). Thousand Oaks, CA: Sage.

Roskos-Ewoldsen, B., Roskos-Ewoldsen, D. R., Yang, M., & Lee, M. (2007). Comprehension of the media. In D. R. Rosks-Ewoldsen & J. Monahan (Eds.), *Communication and social cognition: Theories and methods*. Mahwah, NJ: Lawrence Erlbaum.

Slater, M. D. (2002). Entertainment education and the persuasive impact of narratives. In M. C. Green, J. J. Strange & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 157–181). Mahwah, NJ: Lawrence Erlbaum.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, **21**, 255–287. doi:10.1080/01638539609544959

Tzeng, Y. (2007). Memory for narrative texts: How do parts of the Landscape Model work? *Chinese Journal of Psychology*, **49**, 245–269.

van den Broek, P., & Gustafson, M. (1990). Comprehension and memory for texts: Three generations of reading research. In S. R. Goldman, A. C. Graesser & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence* (pp. 15–34). Mahwah, NJ: Lawrence Erlbaum.

van den Broek, P., Kendeou, P., Sung, Y. C., & Chen, M. (2003). *Comprehension and memory of science texts*: *A simulation using the Landscape Model*. Paper presented at the Annual Meeting of the Society for Text and Discourse, Madrid, Spain.

van den Broek, P., Risden, K., Fletcher, C., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 165–188). Mahwah, NJ: Lawrence Erlbaum.

van den Broek, P., Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch Jr., E. J. O'Brien & R. F. Lorch (Eds.), *Sources of coherence in reading* (pp. 353–373). Hillsdale, NJ: Lawrence Erlbaum.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Mahwah, NJ: Lawrence Erlbaum.

Wyer, R. S. (2004). *Social comprehension and judgment: The role of situation models, narratives, and implicit theories*. Mahwah, NJ: Lawrence Erlbaum.

Yang, M., & Roskos-Ewoldsen, D. R. (2007). The effectiveness of brand placements in the movies: Levels of placements, explicit and implicit memory, and brand-choice behavior. *Journal of Communication*, **57**, 469–489. doi:10.1111/j.1460-2466.2007.00353.x

Zwaan, R. A. (1999). Five dimensions of narrative comprehension: The event-indexing model. In S. R. Goldman, A. C. Graesser & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence* (pp. 93–110). Mahwah, NJ: Lawrence Erlbaum.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, **6**, 292–297. doi:10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 386–397. doi:10.1037/0278-7393.21.2.386

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, **123**, 162–185. doi:10.1037/0033-2909.123.2.162

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, **2**, 199–220. doi:10.1207/s1532799xssr0203_2