

Article

Encoding Systems and Evolved Message Processing: Pictures Enable Action, Words Enable Thinking

Annie Lang ^{1,*}, Rachel L. Bailey ² and Sean Ryan Connolly ¹

¹ Department of Telecommunications, Indiana University, Bloomington, IN 47405, USA; E-Mails: anlang@indiana.edu (A.L.), seanconno@indiana.edu (S.R.C.)

² School of Communication, Washington State University, Pullman, WA 99164, USA; E-Mail: rachel.bailey@wsu.edu (R.L.B.)

* Corresponding author

Submitted: 13 February 2015 | In Revised Form: 28 August 2015 | Accepted: 14 September 2015 |
Published: 28 September 2015

Abstract

This paper, based on theories of ecological perception, embodied motivated cognition, and evolutionary psychology, proposes that pictures elicit evolved biologically imperative responses more quickly and thoroughly than do words. These biologically imperative responses are directly responsible for evolved automatic reactions away from biological threats (e.g. escaping predators, avoiding disease and noxious stimuli) and towards opportunities (e.g. consuming food, approaching mates, finding shelter) in the environment. When elicited, these responses take time to occur and may delay or interfere with other types of behavior. Thus, when environmental information is presented in pictures (which should elicit larger biological responses than words) biological responses should interfere more with higher order tasks like information processing and cognitive decision-making. To test this proposition we designed an experiment in which participants performed speeded categorizations of 60 pairs of matched pleasant and unpleasant environmental opportunities and threats. They categorized the items based on their form (is this a word or a picture?) or based on how the picture made them feel (is this pleasant or unpleasant to you?). If pictures do elicit greater biologically imperative responses than their word counterparts, participants should be able to make form decisions faster than feeling decisions, especially when presented with words rather than pictures and especially when the words and pictures have less biological relevance. This main proposition was supported. Implications for this proposition in terms of communication theory are discussed.

Keywords

embodied; embedded; dynamic systems; message processing; pictures; words

Issue

This article is part of a regular issue of *Media and Communication*, edited by Professor Bradley Greenberg (Michigan State University, USA) and Professor Elisabeth Klaus (University of Salzburg, Austria).

© 2015 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Over the years, communication researchers have spent a lot of time investigating the differential effects of pictures and words on people's cognition, emotion, memory, and comprehension. Research in this area is done in an attempt to ascertain whether pictures or words are more effective for presenting various types

of information in different kinds of messages. This research has demonstrated that the processing of words and pictures varies along a continuum from abstract to concrete. Here we are concerned only with the processing of concrete pictures and words; that is, pictures and words that stand for or depict real things in the world. Certainly, in the era of the internet this kind of research is more relevant to communication research

than ever. In general, we find that pictures tend to be more arousing, more memorable, more easily processed, and elicit more emotional response than words (Houts, Doak, Doak, & Loscalzo, 2006; Kensinger & Schacter, 2006; Lang, Potter, & Bolls, 1999).

Multiple explanations for these results exist, including: first, that pictures are processed in parallel while words are processed serially, or; second, that pictures are processed automatically while the processing of words is more effortful. We, based on theories of ecological perception (i.e. perception action theory), embodied motivated cognition, and evolutionary psychology, propose instead that the faster and more vigorous responses elicited by pictures compared to words arise because pictures automatically elicit more vigorous or thoroughly expressed evolved biologically imperative responses.

Dynamic human centered communication systems theory (DHCCST) (Lang, 2014) is used as the framework for explaining and testing this theoretical hypothesis for why pictures, compared to words, tend to be processed faster and elicit larger emotional responses. Unlike most communication theories framed within the dominant paradigmatic approach consisting of the linear causal model, sensation-perception theory, and information processing approaches (Lang, 2013) DHCCST is based on a new approach based on the dynamic systems non-linear causal model, perception action theory, and motivated cognition.

DHCCST conceptualizes communication as an evolved, embedded, dynamic adaptive process peculiar to animals acting in the world (Clark, 2008; Gibson, 2013; Sherry, 2011). Therefore, we analyze the differences between pictures and words by considering them from the perspective of an evolved human acting in an environment over-time. Time is essential in this analysis and provides guidance for our theorizing at several scales (Berthenthal, 2007). Time is considered at the evolutionary scale (millions of years), at the historical scale (thousands of years), the behavioral scale (days and hours), and at a cognitive/neurological scale (minutes to milliseconds).

2. From an Evolutionary Timescale

The primary assumption of DHCCST is that humans are animals and that they evolved in this particular world. As a result, we evolved in such a way as to be tuned to perceiving and acting in this world with its combination of forces and media. DHCCST's ecological approach is based on Gibson's ecological or direct perception theory. Sometimes called perception action theory, it is based on the notion that humans evolved in this particular world, which provided us with an environment where air, light, water, and gravity are the media that afford animal life and animate behavior/action. In addition to these life supporting and perception-enabling

media through which we move, the world also contains substances, surfaces, detachable objects, and animals. Surfaces vary in opacity, hardness, texture, etc. Surfaces can be open (ground, walls) or closed (rooms, caves). Surfaces can contain substances. Objects are closed surfaces which can be detached and moved from one place to another. Animals are detachable objects with a closed body envelope who can change their external appearance and display animate behavior.

According to ecological perception theory, the perceptual organs are evolved machines that directly pick up information from the world. The movement (i.e. action) of animals and their perceptual organs through the world results in the direct pickup of information in the world allowing animals to move, find food, avoid danger, and procreate. The sensations of sight, hearing or smell are secondary to the actual pick up of the information. Action, not sensation, mediates perception.

DHCCST defines communication as the transfer of information from one animal to another in an environment over time. Communication is both evolved and adaptive. As evolved animals, humans have a variety of evolved communication encoding systems. DHCCST defines evolved communication encoding systems as those actions or emissions of animals that can be directly perceived by the perceptual system of another animal. For example, marking behavior (i.e. leaving behind a smell through an emission or the behavior of rubbing against something) is an evolved encoding system, and the mark is a form of communication that can be directly perceived by the olfactory system of an animal. Similarly, actions and expressions can be directly perceived by the visual system. Sounds (e.g. screams, moans, sighs, cries, growls, purrs) can be directly perceived by the auditory perceptual system. Animals communicate through a variety of evolved perceptual systems. Bees communicate the location of particularly good nectar through the waggle dance (Riley, Greggers, Smith, Reynolds, & Menzel, 2005), ants secrete scent trails to guide one another towards food (Wilson, 1962). Humans and most mammals can smell your fear (de Groot, Semin, & Smeets, 2014), hear your cries or laughter, see your facial expressions and your actions, and express tenderness and love through touch. These are our evolved communication systems.

The evolved human walking through the world in evolutionary time only encountered substances, surfaces, objects and animals. Their perceptual systems directly picked up, over time, both the invariant aspects of their environment (i.e. things that were not changing as they moved through the world) and its variants (i.e. things that were changing as they moved through the world). Invariants include the shape of a mountain, the texture and orientation of the ground, and the location of trees. As the animal moves, some invariants appear while others are occluded but they remain invariant. Other aspects of the world are vari-

ant. Leaves blow in the wind or fall. Animals appear, move, and disappear. The light changes as clouds move across the sun, or the sun rises and sets. Evolved communication information, whether it is an emission, an action, an expression, a sound, or a touch is variant and is directly perceived. Thus, information passing between animals through evolved encoding systems is directly perceived variant information in the environment and is the basis of communication. The evolved human walking through the world in evolutionary time only encountered substances, surfaces, objects, and animals. They never encountered a picture or a word as these did not yet exist.

3. From a Historical Timescale

And then someone made a drawing--in the sand, in the mud, on a cave wall, with chalk, with paint...and there was a picture. What is a picture? According to Gibson (2013) a picture is a surface that captures the invariant aspects of a place or an object at a time and freezes it. Ecological perception mediated by action includes both variant and invariant information. A picture is not an evolved encoding system. It is a man-made invention that captures the invariant aspects of a thing at a specific moment in time. In particular, it captures shape, color and form. What it doesn't capture are the variants of the object or the place. Thus, a picture of a cat has the form and the color and the shape of a cat but it doesn't have locomotion, sniffing, licking, being ruffled by the wind, or being soft when you touch it. When we see a picture of a cat, we directly perceive the invariants of the cat. But the picture of a cat affords only what is afforded by the visual invariant aspects of a cat provided in the ambient light array (i.e. information available to the visual perceptual system). As a result, the picture affords only looking at a cat, not stroking it, smelling it, or hearing it purr. But the form, shape, and color of the cat are still available for direct perception. DHCCST defines pictures as a manmade representational encoding system. Manmade representational systems are those that contain information that is directly perceivable by some perceptual system, but do not contain all of the perceptual information that would be available in the real object, substance, surface, or animal.

And then there were words. After pictures came words. Words are, in a sense, pictures. They are surfaces that provide invariant edges arranged in various orientations fixed in time. Their shape, color, and form can be directly perceived but these shapes and forms are not related to the shapes and forms of the things that they stand for. The shape of the word cat does not provide any directly perceived information about the invariants of the cat. The only directly perceivable information in a word is the shape, color, and form of the word. Thus, the meaning of the word is not directly perceivable.

4. From the Neurological and Cognitive Timescales

What happens, then, when we see a word? How do we get from directly perceiving the form of the word cat, to knowing that the word "cat" stands for cat. As complex evolved animals, humans consist of multiple nested complex systems. DHCCST pays particular attention to the cognitive/neurological system embedded in a physiological system, embedded in a motivational system, embedded in a perceptual system, embedded in a biological system. Each of these systems evolved to function in this world and constrains, interacts with, facilitates, and inhibits actions and reactions in the other systems over time. These systems are nested in the human both biologically and within time. The neurological system is in a faster band than the cognitive (Berthenthal, 2007). Complex interactions of these systems allow us to act, perceive, feel, and think in the world. DHCCST builds on Hofstadter's (Hofstadter, 2000) conceptualization of an active symbol to bridge the gap between direct perception of the form of the word "cat" and the sensation of the meaning of the word "cat". According to Hofstadter, an active symbol is something in the world which initially elicits neuronal activation associated with its form, but then immediately activates additional neuronal activity associated with its meaning and our knowledge of the thing that the symbol stands for. This activity eventually leads to the cognitive sensation of meaning. Thus, when looking at pictures we directly perceive the form of the thing that is represented and respond to it without sensation. When looking at words we directly perceive the form of the word but not the form of the thing that is symbolized. The form triggers neurological activity that both leads to an understanding of the meaning of the word and triggers the appropriate biological response. Thus, the perception of the form of a word is mediated by action, but the perception of the meaning of a word is mediated by sensation. The perception action neurological loop has been shown to be much shorter and therefore much faster than the perception action loop (Pearson, 1985). Thus, while the form of the word is directly perceived as quickly as the form of the picture, the form of the word carries no evolved invariants that will trigger a biologically imperative response while the form of the directly perceived concrete picture does. The meaning of a word arrives more slowly with completion of the perception sensation loop and triggers a biologically imperative response, should one be appropriate.

DHCCST argues that this difference between a representation (e.g. picture) of something and a symbol (e.g. word) of something is at the heart of the differential responses we see in people to concrete words and pictures. People directly perceive the meaning of a representational picture which leads directly to the evolved biologically imperative response. People do not directly perceive the meaning of a word and there-

fore their initial response to a word is neurological and then cognitive. The faster neurological component of the directly perceived picture automatically triggers the evolved biological response before the sensation of meaning but more slowly than the directly elicited response to the picture.

Having made this claim it makes sense to temper it slightly. Because communication is adaptive, it makes sense that a shouted word (Fire!) should elicit almost the same evolved response as the presence of fire. If words completely removed the ability to act swiftly, without the sensation of meaning and conscious thought, to protect ourselves from threat or seize opportunity, they would not be a particularly adaptive form of communication. Yet we know from many areas of priming research that, for practiced readers, people's physical and biological responses to words do begin before they are consciously perceived and identified. Similarly, priming research has also shown larger responses for pictures than for words (Carr, McCauley, Sperber, & Parmelee, 1982). However, once cognition catches up, we also have the ability to either facilitate or inhibit that evolved response. In this nested system approach it is likely that the repeated pairing of the meaning of the word with the invariant shape of the word results in the shape of the word developing motivational relevance, but this response is smaller than that elicited by the evolved invariants of the shape of a cat.

5. From a Behavioral Timescale

DHCCST sees human behavior as the dynamic self-organization of the human nested systems. At this timescale as the human animal acts in the evolved world it directly picks up information about the surrounding environment and its contents, perceiving what the environment affords to promote survival and task completion.

This direct perception of the information in the world is nested in our biological motivational systems which evolved to promote survival. Hence, danger, threats, food, and mates are directly perceived which leads to biological and motivational reflexes that support protective or approach behavior. Action towards and away from motivationally relevant things in the environment, that is things which afford threat or opportunity, begins with direct perception of those things and the elicitation of biological motivational responses. DHCCST conceptualizes motivational responses using Cacioppo's dual motivational system model (Cacioppo & Berntson, 1994, 1999; Cacioppo & Gardner, 1999; Cacioppo, Gardner, & Berntson, 1997). This model posits two independent motivational systems, the approach or appetitive system and the aversive or defensive system. The two systems have been shown to have different patterns of activation. Of importance here is that the aversive system has been shown to activate

more quickly and more vigorously than the appetitive system. This characteristic of the aversive motivational activation function is called the negativity bias. Initial activation in the direct perceptual, biological (e.g. motor), and motivational systems has been shown to be faster than neurological and cognitive responses to the environment and its contents. For example, Whelan, Hiebert and Pearson (1995) showed that changes in gait (biological) in response to direct perception of variations in the ground occurred before the information could have travelled from the foot to the brain and back again, given the speed of nerve propagation, as would have to happen if the gait change resulted from sensation based perception. Similarly, motivation based alteration in probe stimuli can be seen as early as 8 msec post probe onset. The motivational response can modulate the auditory nerve response at 8 msec post probe onset suggesting that the motivational system is faster than one of the fastest parts of the neurological system (Davis, 1997). Cognitive facilitation or inhibition of these automatic/reflexive actions comes somewhat later. This means that motivationally relevant information in the environment directly elicits biologically imperative responses. Its mere presence begins a response that may or may not be stoppable. If the response is very swift (e.g. a startle response (protective) or an orienting response (approach) they can't be stopped. If it's a slower acting response (e.g. moving forward or back, hiding, etc.) its continuation may be slowed by cognitive inhibition or sped up by cognitive facilitation.

6. Media as Part of the World

DHCCST argues that when an animal encounters motivationally relevant contents in the world they respond in an evolved biologically imperative manner. But what happens when motivationally relevant representations (e.g. pictures) and symbols (e.g. words) are encountered in the environment? Decades of communication research have supported the conclusion that responses to mediated information are very similar, at least at the neurological time scale, to responses to similar information in the world (Reeves & Nass, 1996). It has been found, in previous studies examining the processing of words and pictures, that pleasant and unpleasant words and pictures automatically activate the motivational systems (Lang, 2009; Lang & Yeghyan, 2009; Lee & Lang, 2009) with pleasant words and pictures eliciting appetitive activation and unpleasant words and pictures eliciting aversive activation (P. J. Lang, Greenwald, Bradley, & Hamm, 1993). Thus, both pictures and words elicit motivational activation but, due to the negativity bias, unpleasant pictures and words elicit more activation than pleasant pictures and words.

However, as discussed previously, we are arguing that when perceiving pictures, the biologically evolved

response, either aversive or appetitive, will begin immediately because the form of the motivationally relevant thing can be directly picked up from a picture. However, as time continues and there are no variants to accompany the invariants in the picture, the response may unfold in a less vigorous manner since there are no indications that the picture contains further motivationally relevant information (e.g. animate behavior). On the other hand, when a motivationally relevant word is encountered, there is no direct pickup of the form of the object, only the form of the word. Instead, the neurological response occurs identifying the shape of the word, which then triggers the biological response of identifying threat or opportunity and eventually the cognitive identification of the thing. For this reason, DHCCST argues that symbols (i.e. words) result in a slower, less vigorous, and more modifiable biologically imperative response and then enable its inhibition or facilitation.

These biologically imperative motivational responses happen automatically and quickly in order to support behavior towards opportunity and away from threat. Therefore, they can interfere with other ongoing behaviors. Research demonstrates that automatically elicited motivational responses interfere with and slow learned and ongoing behaviors (Boysen, Berntson, Hannan, & Cacioppo, 1996) and that the larger the biological imperative, the greater the interference with the task.

7. Processing Pictures and Words

What does this mean for the processing of pictures and words? The theory described above offers four basic propositions from which to predict human processing of words and pictures. First, pictures elicit bigger and faster biologically imperative responses than do words. Second, unpleasant things elicit larger biologically imperative responses than do pleasant things. Third, biologically imperative responses happen automatically, take time to occur, and affect not only the timing but also the trajectories of other behavior. They may speed certain types of related behavior (e.g. approaching, avoiding), but may interfere with and slow other behavior (e.g. accessing stored information, making decisions). Fourth, biologically imperative responses can be actively inhibited over time and that inhibition will be more successful during the delayed responses triggered by neuronal activation than during the direct activation of the response by pictures.

To test these propositions we have designed an experiment in which people view, on a computer screen, 60 pairs of matched motivationally relevant concrete pictures and words (e.g. the word apple, a picture of an apple). Half of the picture/word pairs are motivationally relevant and pleasant (e.g. food, sex, babies) and half are motivationally relevant and unpleasant (e.g. blood,

death, weapons, snakes). Participants view and categorize the 120 pictures two times. Once, participants categorize as fast as they can, whether the thing on screen is a picture (representation) or a word (symbol) (called form categorization). The other time, they categorize whether the thing on screen is pleasant or unpleasant (called emotional categorization).

Applying the propositions, we expect slower categorization when biological imperatives are larger. Therefore, in general, (H1) pictures and (H2) negative things should have larger biologically imperative responses and be categorized more slowly than words and positive things.

The next prediction (H3) is that emotional categorization will be slower than form (word/picture) categorization. This is because neither the meaning nor the emotional valence of the word or picture is necessary to categorize whether it is a picture or a word. Therefore, the human system will self-organize, during this task, in such a way as to perform the task most successfully and efficiently. This means that people will automatically (and perhaps even intentionally) inhibit their motivational systems in order to decrease the interference of the biologically imperative response with the encoding system categorization task. This should reduce the size of the response and increase the speed of categorization. However, during the emotional categorization task, it is the activation of the motivational systems in response to the directly perceived stimulus that leads to the sensation of pleasantness or unpleasantness. In order to do this task efficiently participants' need the information provided by the biologically imperative response. Therefore, they will not inhibit the motivational systems (to the same degree), resulting in larger responses and slower categorization.

In addition, it is possible that there may be some interactions. For example, the difference in categorization speed between unpleasant pictures and unpleasant words will likely be larger than the difference between positive pictures and positive words. Research question 1 asks if there are any interactions.

8. Method

8.1. Experimental Design

A 2 (Task: Emotional categorization, Form categorization) × 2 (Encoding System: representation/picture, symbol/word) × 2 (Valence: Pleasant, Unpleasant) × 15 (pictures) fully within-subjects factorial design was used.

8.2. Independent Variables

Stimulus. The stimulus consisted of 60 words and 60 pictures which were representations of the words. The 60 words were concrete words selected from the Affective Norms for English Words (ANEW) system (Brad-

ley & Lang, 1999). The ANEW system contains over 1000 words that are normed for valence and arousal. Thirty of the 60 words chosen were rated as pleasant and 30 were rated as unpleasant. The 60 pictures were selected from Google Images to represent the words. The images were selected controlling for background and size. The focal object (which represented the word) was large and centrally foregrounded and backgrounds were simple and emotionally neutral.

Categorization Task. Participants completed two timed categorizations of the stimulus set. The form categorization task required participants to view the words and pictures in random order and indicate whether each was a word or a picture by pressing one of two indicated keys on the keyboard as quickly as possible. The emotional categorization task used the same procedure but participants were asked to categorize whether the word or picture was emotionally pleasant or unpleasant. Participants were instructed to use fingers on different hands. They used either the R-Shift (positive) versus L-Shift (negative) for the emotion categorization or R-Ctrl (picture) vs. L-Ctrl (word), which are on opposite sides of the keyboard to complete the categorizations. The more natural position for using these keys was two hands, and this was the demonstrated positioning. The keys representing a category were the same for all participants. This should not be a problem as Woods, Wyma, Yund, Herron & Reed (2015) demonstrated that simple categorizations were not affected by handedness except in cases of lateral presentation. In this study there was no lateralization of stimuli. Pictures and words were always presented in the center of the screen.

8.3. Dependent Variables

Response Latency. Items were evaluated one at a time. Individuals were instructed to keep their fingers on the response keys and answer as quickly as possible without sacrificing accuracy. The time of evaluation for each image or word was collected in milliseconds from the time the item appeared to the time the participant pressed a response key. The accuracy of the categorization of each word and picture in both tasks (word/picture or pleasant/unpleasant) was monitored. Proportion correct was computed. Accuracy was slightly higher for the form task (97%) than the emotional valence task (89%).

8.4. Data Analysis

Measures were taken to deal with outliers in these latency data. For each categorization, reaction time outliers over and under 2.49 standard deviations from the mean were replaced with the value of that limit (Cousineau, & Chartier, 2010). In order to do this, Z-scores were calculated. Of the total 11,520 reaction

time responses in both tasks, only 383 (3%) were replaced (187 in the emotional task and 196 in the form task). After replacing the outliers, the Z-scores were then calculated back into milliseconds for analysis. These data were then analyzed as raw data submitted to a 2 (Categorization Task: Emotional, Form) \times 2 (Encoding System: Word, Picture) \times 2 (Valence: Pleasant, Unpleasant) \times 15 (Repetitions) repeated-measures ANOVA.

8.5. Participants

Participants (N = 48) were undergraduates in a telecommunications class and received extra credit in the course for their participation. 60% were females with a median age of 18.

8.6. Procedure

Participants completed the protocol individually. Upon arrival informed consent was obtained after all participants' questions were answered satisfactorily. Then, participants were seated at a laptop station equipped with MediaLab and DirectRT software (Jarvis, 2010), which were responsible for delivering the stimuli and collecting responses. Participants first engaged in an unrelated picture evaluation task. Next, participants completed the emotional and form categorization tasks in random order. Upon completion of these tasks, the participants completed a series of personality scales. Then, participants were thanked, debriefed and dismissed.

9. Results

Hypothesis one predicted that people would categorize pictures as pictures more slowly than they would categorize words as words. The main effect of encoding system was significant, $F(1,47)=17.64$, $p<.00$, $\eta^2=.27$. The average categorization speed for pictures was slower ($M=706.10$, $SE=14.243$) than that found for words ($M=668.40$, $SE=11.40$) across both categorization tasks.

Hypothesis two predicted that classifying feelings would be slower than classifying form. The main effect of task was significant, $F(1,48)=279.82$, $p<.00$, $\eta^2=.86$. On average people categorized how they felt much more slowly ($M=869.40$, $SE=20.47$) than form ($M=505.10$, $SE=10.57$).

Hypothesis three predicted that unpleasant pictures and words would be categorized more slowly than pleasant pictures and words. The main effect of valence only approached significance, $F(1,48)=2.62$, $p<.11$, $\eta^2=.05$, though it was in the correct direction. On average people categorized negative words and pictures more slowly ($M=695.99$, $SE=15.29$) than positive pictures and words ($M=678.51$, $SE=10.82$).

Research question one asked about the possibility of interactions. There were no significant interactions, though the task \times valence interaction approached sig-

nificance, $F(1,47) = 1.84$, $p < .18$, $\eta^2 = .04$. During form categorization (when the motivational systems are inhibited) there is virtually no difference in the speed of categorizing words ($M=503.31$, $SE=10.00$) and pictures ($M=506.90$, $SE=11.46$). During valence categorization, however, pictures are categorized more slowly ($M=885.09$, $SE=26.89$) than words ($M=853.71$, $SE=18.07$), $t(47)=2.97$, $p < .01$.

10. Discussion

This paper uses DHCCST to reconceptualize communication from an evolved, embedded, embodied, dynamic perspective. The difference between pictures and words is considered to be a difference in perception. Pictures are thought to enable direct perception (or perception mediated by action) while words are thought to allow direct perception only of the shape of the word. Subsequent neurological and cognitive activity is required before the meaning of a word is understood (sensation mediated perception). For this reason it is argued that pictures elicit faster and more intense biologically imperative responses (associated with approach and avoidance), which interfere with ongoing tasks. Thus, pictures, through direct perception, elicit action, especially that action which is in service of biologically imperative behaviors (consuming food, finding mates, avoiding danger). Words on the other hand, elicit a delayed biological response which can then be inhibited or facilitated once it is activated. For this reason, words allow us to think about motivationally relevant material in a more cognitive and less motivational fashion.

Using this perspective it was predicted that 1) people would categorize pictures as pictures more slowly than they would categorize words as words; 2) that they would perform form categorization faster than emotional categorization, and; 3) that they would categorize negative words and pictures more slowly than positive words and pictures. Hypothesis 1 and 2 were supported. Hypothesis 3 was partially supported by the nearly significant task X valence interaction which showed no difference between pleasant and unpleasant pictures and words during form categorization but a significantly slower categorization of unpleasant compared to pleasant pictures and words during emotional categorization.

The primary importance of these findings may be that they provide initial support for the approach being advocated by DHCCST but there are also some implications for communication research. A great deal of communication research has focused on the use of pictures compared to text or television compared to print. In general pictures are found to be remembered somewhat better than text (Finnegan & Vishwanath, 1996), to elicit somewhat more attention (Houts et al., 2006), and to improve memory for text, especially when the topic of the picture or the news is concrete

(David, 1998). In addition, it has been suggested that picture processing requires fewer cognitive resources compared to the processing of words (Lang et al., 1999). And while the evidence is somewhat mixed it has been suggested that pictures increase comprehension, attention, and memory even more for people with lower levels of literacy (Finnegan & Vishwanath, 1996; Houts et al., 2006).

The results of this study support the contention that concrete pictures, because they are directly perceived, activate very fast acting biological and motivational systems, which then slave the initial response of slower acting, higher order systems. This direct perception is not related to literacy or education or previous knowledge, but only to our nature as evolved humans. As a result, pictures, and in particular those that are threats or opportunities (i.e. motivationally relevant), will be perceived, attended to, and understood by these lower acting processes faster and at a lower level than the same information presented in words. This could be the underlying mechanism for the higher order cognitive effects described above.

It is also worth pointing out that the results of this study do not correspond to the predictions of more cognitive, information processing based approaches to the processing of words and pictures and standard interpretations of simple reaction time as it relates to processing, resource allocation, and attention. For example, the LC4MP argues that motivationally relevant compared to non-motivationally relevant media content both elicits a greater allocation of cognitive resources and requires their use. As a result, as motivational relevance increases motivationally relevant material is thought to be harder to process and to lead more quickly to cognitive overload. This prediction has been supported using secondary task reaction times and recognition as a combined indicator of available resource (i.e. resources allocated–resources required) both for audio (verbal/word) and visual (picture) information. To the extent that pictures and words are processed differently, the LC4MP argues that pictures may improve memory because they require fewer resources to process than words and therefore encounter cognitive overload less quickly compared to verbal and text stimuli. But these arguments of automatic processing and fewer resources for pictures do not translate into a prediction that pictures will be categorized as pictures more slowly than words would be categorized as words. Rather the more likely prediction would be that motivationally relevant information (whether a picture or a word) should be categorized more quickly than non-motivationally relevant information and that effect would be larger for negative motivationally relevant information due to the negativity bias. But the results of this study are the opposite of that prediction. Instead negative motivationally relevant images are categorized as images more slowly than any of the other categories.

The results of this study are in line with communication research that theorizes that negative information (and in these studies primarily pictorial or audio/visual negative information) is compelling (Grabe & Kamhawi, 2006; Newhagen & Reeves, 1992; Shoemaker, 1996). This research is based on the idea that we evolved to keep track of threats in the environment and therefore that we automatically attend to negative information. Again, the suggestion here is that motivationally relevant pictures are indeed directly perceived and that the elicited biological imperative (which evolved to protect the animal) precedes and interferes with other tasks and processes. In other words, motivationally relevant pictures, in particular those of primary motivators such as sex, food, and danger, should interfere more with other tasks, thereby compelling the processing of the motivationally relevant material.

Finally, these results provide encouraging support for the DHCCST perspective. Future research in this area should be done to attempt to track the underlying processes theorized to result in the behavioral results found in this paper. For example, one idea would be to include neutral pictures and words as a comparison. DHCCST would predict that neutral words and pictures depicting objects would not differ in the time it takes to categorize form and valence (as pleasant, unpleasant, or neutral) because they would not elicit biologically imperative responses to delay and interfere with the task.

Building on direct perception research, future research could compare word/picture pairs of objects with those of pleasant and unpleasant animals. Given that animals are more consequential than objects (because they can engage in animate behavior) one might expect them to elicit larger biological responses than objects. Building on theories of motivational relevance, future studies might also compare word/picture pairs that differ in their levels of motivational relevance. Some pairs might represent primary reinforcers (food, sex, blood, attacking animals) and others things that are merely pleasant and unpleasant (flowers, garbage, etc.) Here the prediction would be that the primary reinforcer word/picture pairs would have a larger biological imperative than the merely pleasant and unpleasant word/picture pairs. Building on the suggestion made earlier, that the shape of a word may become motivationally relevant over repeated pairings, might be tested by examining categorization speed differences in pairs of highly arousing and calm words of different frequency of occurrence in the English language. Finally, future research might use neurological and physiological indicators of motivational activation to ascertain if indeed emotional pictures lead to faster neurological and motivational responses compared to words and if motivational activation is lower during form compared to picture categorization. This would provide additional support for the notion that timed response data can be

used as an indicator of the level of inhibition of the motivational systems.

Conflict of Interests

The authors declare no conflict of interests.

References

- Bertenthal, B.I. (2007). Dynamical systems: It's about time! In S. Boker (Ed.), *Data analytic techniques for dynamical systems*. Hillsdale, NJ: Erlbaum.
- Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, 22(1), 76-86.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW)*. Gainesville, FL: The NIMH Center for the Study of Emotion and Attention, University of Florida.
- Cacioppo, J. T., & Bernston, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of pleasant and unpleasant substrates. *Psychological Bulletin*, 115(3), 401-423.
- Cacioppo, J. T., & Berntson, G. G. (1999). The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science*, 8(5), 133-137.
- Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual Reviews: Psychology*, 50, 191-214.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1, 3-25.
- Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 8(6), 757.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment: A review. *International Journal of Psychological Research*, 3(1), 58-67.
- David, P. (1998). News concreteness and visual-verbal association: Do news pictures narrow the recall gap between concrete and abstract news? *Human Communication Research*, 25(2), 180-201.
- Davis, M. (1997). The neurophysiological basis of acoustic startle modulation: Research on fear motivation and sensory gating. In *Attention and orienting: Sensory and motivational processes* (pp. 69-96). New York: Psychology Press.

- de Groot, J. H. B., Semin, G. R., & Smeets, M. A. M. (2014). I can see, hear, and smell your fear: Comparing olfactory and audiovisual media in fear communication. *Journal of Experimental Psychology: General*, *143*(2), 825-834. doi:http://dx.doi.org/10.1037/a0033731
- Finnegan, J. R., & Viswanath, K. (1996). The knowledge gap hypothesis: Twenty five years later. In *Communication yearbook* (pp. 117-135). Thousand Oaks, CA: Sage.
- Gibson, J. J. (2013). *The ecological approach to visual perception*. New York: Psychology Press.
- Grabe, M. E., & Kamhawi, R. (2006). Hard wired for negative news? Gender differences in processing broadcast news. *Communication Research*, *33*(5), 346-369.
- Hofstadter, D. R. (2000). *Godel, Escher, Bach*: New York: Penguin.
- Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J. (2006). The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling*, *61*(2), 173-190.
- Jarvis, B. G. (2010). *MediaLab research software*. New York: Empirisoft.
- Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(2), 110-126.
- Lang, A. (2013). Discipline in crisis? The shifting paradigm of mass communication research. *Communication Theory*, *23*, 10-24. doi:10.1111/comt.12000
- Lang, A. (2014). Dynamic human-centered communication systems theory. *Information Society*, *30*(1), 60-70.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261-273. doi:10.1111/j.1469-8986.1993.tb03352
- Lang, A. (2009). The limited capacity model of motivated mediated message processing. In R. Nabi & M. B. Oliver (Eds.), *The Sage handbook of mass media effects* (pp. 193-204). Thousand Oaks, CA: Sage.
- Lang, A., Potter, R. F., & Bolls, P. D. (1999). Something for nothing: Is visual encoding automatic? *Media Psychology*, *1*(2), 145-163.
- Lang, A., & Yeghyan (2009). Motivated message processing: How media elicit motivation which influences how media are processed. In J. McCroskey, K. Floyd, & M. Beatty (Eds.), *Biological dimensions of communication* (pp. 135-159). New York: Hampton Press.
- Lee, S., & Lang, A. (2009). Discrete emotion and motivation: Relative activation in appetitive and aversive motivational system as a function of anger, sadness, fear, and joy during televised information campaigns. *Media Psychology*, *12*, 148-170.
- Newhagen, J. E., & Reeves, B. (1992). The evening's bad news: Effects of compelling negative television news images on memory. *Journal of Communication*, *42*(2), 25-41.
- Pearson, K. G. (1985). Are there central pattern generators for walking and flight in insects? In *Feedback and motor control in invertebrates and vertebrates* (pp. 307-315). New York: Springer.
- Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television? New media like real people? Places. *Computers and Mathematics with Applications*, *5*(33), 128.
- Riley, J., Greggers, U., Smith, A., Reynolds, D., & Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature*, *435*(7039), 205-207.
- Sherry, J. (2011). *Communication science in the 21st century*. East Lansing: Michigan State University.
- Shoemaker, P. J. (1996). Hardwired for news: Using biological and cultural evolution to explain the surveillance function. *Journal of Communication*, *46*(3), 32-47.
- Whelan, P. J., Hiebert, G. W., & Pearson, K. G. (1995). Plasticity of the extensor group I pathway controlling the stance to swing transition in the cat. *Journal of Neurophysiology*, *74*(6), 2782-2787.
- Wilson, E. O. (1962). Chemical communication among workers of the fire ant *Solenopsis Saevisissima*: The Organization of Mass-Foraging. *Animal Behaviour*, *10*(1), 134-147.
- Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience*, *9*.

About the Authors



Dr. Annie Lang

Annie Lang (Ph.D. University of Wisconsin at Madison) is a Distinguished Professor of Communication Science and Cognitive Science at Indiana University. Her research seeks to explain how people process mediated messages. To date this work has produced a general data-driven model of mediated message processing. This research has resulted in over 100 academic publications and been supported by almost 3 million dollars in grants. She is a Fellow of the International Communication Association and recipient of ICA's Steven H. Chaffee career productivity award.



Dr. Rachel L. Bailey

Rachel L. Bailey (Ph.D. Indiana University) is an assistant professor in The Edward R. Murrow College of Communication at Washington State University. She directs the Communication Emotion and Cognition Lab. Her research focuses on the complex interactions of biological and environmental factors that influence human communication and health-related behaviors.



Sean Ryan Connolly

Sean Connolly (MA Communications, MS Computer-Human-Interactions) is a media researcher. He focuses on the “user experience” of narrative entertainment, with a focus on symbols, myth, and cinematic language.