

Georgia State University

ScholarWorks @ Georgia State University

Psychology Theses

Department of Psychology

10-15-2009

Gaze Fixation during the Perception of Visual and Auditory Affective Cues

Susan M. McManus
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/psych_theses



Part of the [Psychology Commons](#)

Recommended Citation

McManus, Susan M., "Gaze Fixation during the Perception of Visual and Auditory Affective Cues." Thesis, Georgia State University, 2009.

doi: <https://doi.org/10.57709/1331534>

This Thesis is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

Appendix 2

Condition	Variable	Effect	χ^2	df	<i>p</i>	ε
Task 1 Core vs Periphery	Duration	Emotion	12.80	5	.03	.88
Task 1 Core Regions	Duration	Region	29.24	2	.00	.69
	Count	Region	40.75	2	.00	.64
Task 1 Core Regions x Emotions	Duration	Region	21.78	5	.00	.79
		Emotion x Region	84.41	44	.00	.74
	Count	Region	36.22	5	.00	.68
Task 2 Congruent vs Incongruent	Duration	Region	11.40	2	.00	.83
Task 2 Fearful Movies	Duration	Movie x Region	9.63	2	.01	.85
Task 2 Happy Movies	Duration	Region	11.38	2	.00	.83
	Count	Region	17.85	2	.00	.77
Task 2 Neutral	Duration	Region	7.04	2	.03	.88
	Count	Region	12.13	2	.00	.82
		Movie x Region	7.12	2	.03	.88
Task 3 Congruent vs Incongruent	Duration	Movie x Region	8.12	2	.02	.87
	Count	Region	18.46	2	.00	.76
Task 3 Angry Movies	Duration	Region	7.09	2	.03	.88
	Count	Region	20.95	2	.00	.74
Task 3 Fearful Movies	Count	Region	8.04	2	.02	.87
Task 3 Happy Movies	Duration	Region	8.56	2	.01	.86
	Count	Region	20.93	2	.00	.74
Task 3 Neutral Movies	Duration	Region	7.43	2	.02	.88
	Count	Region	21.53	2	.00	.74
		Emotion x Region	9.24	2	.01	.85

Appendix 1

The study will begin shortly... message will appear on the screen → Task Instructions

“I have placed this plastic/gel cover over the keyboard so that only four keys are exposed (*hold the keyboard up for the participant to see*) and those are the keys you will use to respond. After the ‘study...’ message goes away, a reminder of the directions for this task will appear on the screen. You are to pay attention to each video and answer the question, “How does this person feel?” Before each movie begins, a crosshair (plus sign) will appear in the middle of the screen (*point to the center of the screen as you say this*). That is a cue for you to look to that point before each video clip begins. Each video clip will play for approximately 2 seconds, after which four emotion choices will appear on the screen. The choices will be Happy, Angry, Fearful, Neutral” (*point to the positions on the screen as you say this*) and these choices will line up spatially (from left to right) with the keys on the keyboard. You can just rest four fingers on the exposed keyboard keys and then you won’t need to see the keyboard to make your response. As soon as you make your response, it will automatically advance to the next video clip. If you do not respond in the 5-second time allowed by the program, you will see a message that says ‘Please try and respond faster’. Do you have any questions?”

Practice Run:

“Because this is practice, you will only see pictures”

Task 2:

“The instructions for this next task will be exactly the same as before”

Task 3:

“This time, the instructions will be different. Instead of answering the question of how the person feels, you will answer the question, “Does the emotion in the face match the emotion in the voice?” You will respond either Yes or No and those response choices will again come up on the screen. Yes will always be on the left and No will always be on the right. Those responses will correspond with the left and right buttons on the mouse.”

Psychology, 112(3), 371-381.

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.

Yin, R.K. (1970). Face recognition by brain-injured patients: A dissociable ability? *Neuropsychologia*, 8, 395-402.

- Tarr, M.J. & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level Visual processing automatized by expertise. *Nature Neuroscience*, 3(8), 764-769.
- Teunisse, J. P. & de Gelder, B. (2003). Face processing in adolescents with autistic disorder: The inversion and composite effects. *Brain and Cognition*, 52, 285-294.
- Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology*, 17, 257-279.
- Valentine, T. (1988) Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79, 471-491.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective, & Behavioral Neuroscience*, 1(4), 382-387.
- Vuilleumier, P. & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional from functional neuroimaging. *Neuropsychologia*, 45, 174-194.
- Vuilleumier, P., Richardson, M.P., Armony, J.L., Dolan, R.J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7(11), 1271-1278.
- Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, 6, 313-326.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.
- White, M. (1999). Representation of facial expressions of emotion. *American journal of*

- emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38, 1371-1381.
- Schwarzer, G., Huber, S., & Dummler, T. (2005). Gaze behavior in analytical and holistic Face processing. *Memory & Cognition*, 33(2), 344-354.
- Seashore, C., Lewis, D., & Saetveit, J. (1960). *Seashore Rhythm Test*. In: Seashore, *Measures of Musical Talents, Revised Manual*. New York: The Psychological Corporation.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, 75, 221-242.
- Smith, M.L., Cottrell, G.C., Gosselin, F. & Shyns, P.G. (2005) Transmitting and decoding facial expressions. *Psychological Science*, 16(3), 184-189.
- Stacey, P. C., Walker, S. & Underwood, J.D. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology*, 96, 407-422.
- Tamietto, M., Corazzini, L.L., de Gelder, B., & Geminiani, G. (2006). Functional asymmetry and interhemispheric cooperation in the perception of emotions from facial expressions. *Experimental Brain Research*, 171, 389-404.
- Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, 46, 225-245.
- Tanaka, J.W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23(3), 457-482.
- Tantam, D., Monaghan, L., Nicholson, H., & Stirling, J. (1989). Autistic children's ability to interpret faces: a research note. *Journal of Child Psychology and Psychiatry*, 30, 623-630.

- audiovisual speech perception: The influence of ocular fixations on the McGurk effect. *Perception & Psychophysics*, 65(4), 553-567.
- Pell, M.D. (2005). Nonverbal emotion priming: Evidence from the 'facial affect decision task'. *Journal of Nonverbal Behavior*, 29(1), 45-73.
- Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual Scanning of Faces in Autism. *Journal of Autism and Developmental Disorders*, 32(4), 249-261.
- Pourtois, G., de Gelder, B., & Vroomen, J. (2000). The time-course for intermodal binding between seeing and hearing affective information. *Neuroreport: For Rapid Communication of Neuroscience Research*, 11(6), 1329-1333.
- Reynolds, C.R. & Kamphaus, R.W. (2004). BASC-2 Manual. AGS Publishing: Circle Pines, MN.
- Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinates of facial appearance. *Perception*, 17, 43-63.
- Rhodes, G., Brake, S., & Atkinson, A.P. (1993). What's lost in inverted faces? *Cognition*, 47, 25-57.
- Robins, D.L., Hunyadi, E., & Schultz, R.T., (2009). Superior temporal activation in response to dynamic audio-visual emotion cues. *Brain and Cognition*, 69, 269-278.
- Robins, D.L., & Schultz, R.T. (2004). Cross-modal integration of emotional information. Paper presented at the International Neuropsychological Society, February, 2004, Baltimore, MD.
- Rutherford, M.D., & Towns, A.M., (2008). Scan path differences and similarities during

- Mertens, I., Siegmund, H., & Grusser, O. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia*, 31, 989-998.
- Mitchell, R.L., Elliot, E., Barry, M. Cruttenden, A., & Woodruff P.W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, 41, 1410-1421.
- Mitchell, R.L. (2006). How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. *European Journal of Neuroscience*, 24,3611-3618.
- Mondloch, C.J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, 31, 553-566.
- Noton, D. & Stark, L (1971). Scanpaths in eye movements during pattern perception. *Science*, 171, 308-311.
- Nowicki, S. Jr. (2004). A test and manual for the Diagnostic Analysis of Nonverbal Accuracy 2. Unpublished manuscript, Emory University, Atlanta, GA.
- Nowicki, S. & Carton, J. (1993). The measurement of emotional intensity from facial expressions. *The Journal of Social Psychology*, 133(5), 749-750.
- Nowicki, S. & Duke, M.P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of nonverbal Behavior*, 18(1), 9-35.
- Nusseck, M., Cunningham, D.W., Wallraven, C., & Bulthoff, H.H. (2008). The contribution of different facial regions to the recognition of conversational speech. *Journal of Vision*, 8(8), 1-23.
- Pare, M., Richler, R.C., ten Hove, M., Munhall, K.G., (2003). Gaze behavior in

- Lindner, J.L. & Rosen, L.A. (2006). Decoding of emotion through facial expression, prosody, and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, 36(6), 769-777.
- Lobmaier, J.S., Klaver, P., Loenneker, T., Martin, E., & Mast, F.W. (2008). Featural and configural face processing strategies: Evidence from a functional magnetic resonance imaging study. *NeuroReport: For Rapid Communication of Neuroscience Research*, 19(3), 287-291.
- Luria, S.M., & Strauss, M.S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception and Psychophysics*, 2, 547-552.
- Mackworth, N.H., & Bruner, J.S. (1970). How adults and children search and recognize pictures. *Human Development*, 13, 149-177.
- Manor, B.R., Gordon, E., Williams, L.M., Rennie, C.J., Bahramali, H., Latimer, R.J., & Meares, R.A. (1999). Eye movements reflect impaired face processing in patients with schizophrenia. *Society of Biological Psychiatry*, 46, 963-969.
- Marcus, D.J. (2005). *Processing of facial expression information in children and adults: a behavioral and psychophysiological study*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.
- Massaro, D.W. (1998). *Perceiving talking faces: From speech perception to a behavior principle*. Cambridge, MA: MIT Press.
- Massaro, D.W., & Egan, P.B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review*, 3(2), 215-221.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.

- human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302-4311.
- Kanwisher, N. & Moscovitch, M. (2000). The cognitive neuroscience of face processing: an introduction. *Cognitive Neuropsychology*, 17, 1-11.
- Kanwisher, N. & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B*, 361, 2109-2128.
- Kington, J.M., Jones, L.A., Watt, A.A., Hopkin, E.J., & Williams, J. (2000). Impaired eye expression recognition in schizophrenia. *Journal of Psychiatric Research*, 34(4-5), 341-347.
- Kinsbourne, M. (1972). Eye and head turning indicates cerebral lateralization. *Science*, 176(4034), 539-541.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59(9), 809-816.
- Kohler, C.G., Turner, T., Stolar, N.M., Bilker, W.B., Brensinger, C.M., Gur, R.E., & Gur, R.C. (2004). Differences in facial expression of four universal emotions. *Psychiatry Research*, 128, 235-244.
- Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research*, 184(4), 533-546.
- Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, Hearing, and Resolution*, 42, 526-539.

- Hietanen, J., Leppanen, J.M., Illi, M., & Surakka, V., (2004). Evidence for the integration of audiovisual emotional information at the perceptual level of processing. *European Journal of Cognitive Psychology*, 16(6), 769-790.
- Hobson , R.P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: coordinating faces and voices. *Psychological Medicine*, 18, 911-923.
- Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli. *Memory and Cognition*, 4(2), 176-185.
- Horley, K., Williams, L.M., Gonsalvez, C., & Gordon, E. (2003). Social phobics do not see eye to eye: A visual scanpath study of emotional expression processing. *Journal of Anxiety Disorders*, 17(1), 33-44.
- Horley, K., Willlliams, L.M., Gonsalvez, C., & Gordon, E. (2004). Face to face: visual scanpath evidence for abnormal processing facial expressions in social phobia. *Psychiatry Research*, 127(1-2), 43-53.
- Itier, R.J., Alain, C., Sedore, K., & McIntosh, A.R. (2007). Early face processing specificity: It's in the eyes! *Journal of Cognitive Neuroscience*, 19(11), 1815-1826.
- Jarvis, B. (1999). DirectRT Precision Timing Software. Empirisoft Software.
- Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children With autism. *Journal of Child Psychology and Psychiatry*, 44(4), 529-542.
- Kamio, Y., Wolf, J. & Fein, D. (2006). Automatic processing of emotional faces in high functioning pervasive developmental disorders: an affective priming study. *Journal of Autism and Developmental Disorders*, 36(2), 155-167.
- Kanwisher, N., McDermott, & Chun, M. (1997). The fusiform face area: a module in

- for a generalized deficit? *Child Neuropsychology* 2(2), 123-139.
- Graves, C.A. & Natale, M. (1979). The relationship of hemispheric preference, as measured by conjugate lateral eye movements, to accuracy of emotional facial expression. *Motivation and Emotion*, 3(3), 219-234.
- Green, M.J., Williams, L.M., & Davidson, D. (2003). In the face of danger: Specific viewing strategies for facial expressions of threat? *Cognition and Emotion*, 17(5), 779-786.
- Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of gaze behavior research* (pgs. 523-533). Amsterdam: Elsevier.
- Gross, T.F. (2005). Global-local precedence in the perception of facial age and emotional expression by children with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders*, 35(6), 773-785.
- Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal Bias in Recognition of Facial Emotions in Children with Asperger Syndrome. *Journal of Child Psychology and Psychiatry*, 41(3), 369-379.
- Gur, R.E. (1975). Conjugate lateral eye movements as an index of hemispheric activation. *Journal of Personality and Social Psychology*, 31(4), 751-757.
- Haxby, J.V., Hoffman, E.A., & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, 4, 223-233.
- Hecht, D., & Reiner, M. (2008). Sensory dominance in combinations of audio, visual, and haptic stimuli. *Experimental Brain Research*, 193, 307-314.

- Configural and local processing of faces in children with Williams Syndrome. *Brain and Cognition*, 41, 276-298.
- Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107-117.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about Face perception? *Psychological Review*, 105 (3), 482-498.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Fusar-Poli, P., Placentino, A., Carletti, F., Allen, P., Landi, P., Abbamonte, M., Barale, F., Perez, J., McGuire, P., & Ploiti, P.L. (2009). Laerality effect on emotional face processing: ALE meta-analysis of evidence. *Neuroscience Letters*, 452, 262-267.
- Gamer, M., & Buchel, C. (2009). Amygdala activation predicts gaze toward fearful eyes. *The Journal of Neuroscience*, 29(29), 9123-9126.
- Gauthier, I. & Tarr, M.J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1673-1682.
- Gauthier, I., Tarr., M.J., Anderson, A.W., Skudlarski, P., & Gore, J.C. (1999). Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573.
- Gauthier, I., Williams, P. Tarr, M.J., & Tanaka, J. (1998). Training ‘Greeble’ experts: a framework for studying expert object recognition processes. *Vision Research*, 38, 2401-2428.
- Gepner, B., deGelder, B. & deSchonen, S. (1996). Face processing in autistics: evidence

- Buchan, J.N., Pare, M., & Munhall, K.G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, 2(1), 1-13.
- Buchan, J.N., Pare, M., & Munhall, K.G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audiovisual speech perception, *Brain Research*, 1242, 162-171.
- Bukach, C.M., Bub, D.N., Gauthier, I., & Tarr, M.J. (2006). Perceptual expertise effects are not all or none: spatially limited perceptual expertise for faces in a case of prosopagnosia. *Journal of Cognitive Neuroscience*, 18(1), 48-63.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *TRENDS in Cognitive Sciences*, 10(4), 159-166.
- Calvo, M.G. & Lang, P.J. (2004). Gaze patterns when looking at emotional pictures: motivationally biased attention. *Motivation and Emotion*, 28(3), 221-243.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ, Lawrence Erlbaum Associates.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126-135.
- Davidson, R.J. Ekman, P., Saron, C.D., Senulis, J.A., & Friesen, W.V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology 1. *Journal of Personality and Social Psychology*, 58(2), 330-341.
- de Gelder, B., & Vroomen, J. (2000). The Perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- Dereulle, C., Mancini, J., Livet, M.O., Casse-Perrot, C., & de Schonen, S. (1999).

- Behrmann, M., Avidan, G., Marotta, J.J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: behavioral findings. *Journal of Cognitive Neuroscience*, 17(7), 1-19.
- Benton, A. (1994). *Face Recognition*. Los Angeles: Western Psychological Services.
- Bilenko, N.Y., Grindrod, C.M., Myers, E.B., & Blumstein, S.E. (2008). Neural correlates of semantic competition during processing of ambiguous words. *Journal of Cognitive Neuroscience*, 21(5), 960-975.
- Borod, J.C., Vingiano, W., & Cytryn, F. (1988). The effects of emotion and ocular dominance on lateral eye movement. *Neuropsychologia*, 26(2), 213-220.
- Borod, J.C., Cicero, B.A., Obler, L.K., Welkowitz, J., Erhan, H.M., Santschi, C., Grunwald, I.S., Agosti, R.M., & Whalen, J.R. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology*, 12(3), 446-458.
- Bourne, V.J., Vladeanu, M., & Hole, G.J. (2009). Lateralised repetition priming for featurally and configurally manipulated familiar faces: Evidence for differentially lateralized processing mechanisms. *Laterality: Asymmetries of Body, Brain, and Cognition*, 14(3), 287-299.
- Bradley, M.M., & Lang, P.J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavioral Therapy & Experimental Psychiatry*, 25(1), 49-59.
- Bradley, M.M., Miccoli, L., Escrig, M.A., & Lang, P.J., (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45, 602-607.
- Bradshaw, J.L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception & Psychophysics*, 9(5), 443-448.

References

- Adolphs, R., Jansari, A., & Tranel, D. (2001). Hemispheric perception of emotional valence from facial expressions. *Neuropsychology*, 15(4), 516-524.
- Ambadar, Z., Schooler, J.W., & Cohn, J.F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, 16(5), 403-410.
- Asthana, H.S., & Mandal, M.K. (2001). Visual-field bias in the judgment of facial expressions of emotion. *Journal of General Psychology*, 128(1), 21-29.
- Aviezer, H., Hassin, R.R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies in the malleability of emotion perception. *Psychological Science*, 19(7), 724-732.
- Bar-Haim, Y.C., Shulman, C., Lamy, D. & Reuveni, A. (2006). Attention to eyes and mouth in high-functioning children with autism. *Journal of Autism and Developmental Disorders*, 36(1), 131-137.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome, high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5-17.
- Bate, S., Haslam, C., & Hodgson, T.L. (2009). Angry faces are special too: Evidence from the visual scanpath. *Neuropsychology*, 23(5), 658-667.
- Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, 44, 110-129.

tasks (affective judgment, match/mismatch). This would strengthen the argument that there are standard fixation patterns that exist across the four emotions included in this study and that the change in pattern that emerges during incongruent movies is task dependent.

Finally, in order to support the use of the DAVE stimuli in emotion perception tasks, a comparison of processing approaches and fixation patterns elicited during these novel dynamic stimuli should be made with previously validated static photos is suggested. Furthermore, examination of the DAVE stimuli as visual-only stimuli would provide additional understanding of how the visual scanning of dynamic faces differs from that of static faces. These stimuli could then be used to evaluate the differences between emotion perception strategies of typically developing and other populations, such as individuals with autism spectrum disorders, in which emotion recognition deficits have been found. Specifically, it is suggested that the emotion identification abilities and visual scanning patterns of a sample of young adults with autism be explored through the use of the DAVE stimuli. Previous research suggests that individuals with autism are less influenced by facial information, instead showing a bias toward verbal or auditory information, and rely on affective information from the mouth region more than the eyes. The direct comparison of these differences through eye-tracking techniques and during the perception of dynamic audio-visual stimuli would provide further understanding as to where these processes break down and may provide suggestions for intervention strategies.

Task requirements appeared to have affected the visual scanning approach used in the current study. The patterns of fixations that emerged during the congruent movies across both sets of instructions – emotion identification and congruence detection – appeared to be similar, but were not compared directly. However, the presence of an incongruent auditory emotion caused participants to increase fixations to the regions most relevant to the task instructions. Under requirements of emotion judgment, participants were asked to label explicitly the emotion and looked to regions most salient for emotion identification. Participants showed a bias toward the visual affective information, in that they answered with the facial emotion. It therefore is logical that they would increase fixations to the eyes, a region most frequently identified as the most salient visual affective source. In contrast, the match/mismatch task led participants to look for additional affective cues in the face, cues that would confirm the mismatch that was perceived from auditory information.

There are some limitations to the current study that are worth mentioning. The short duration of the movie clips limited the amount of information that could be gathered about processing strategy or fixation pattern that emerges over a period of time. This limitation may be addressed by developing DAVE stimuli of a longer duration. In addition, several inferences were made about the lateralization of emotion perception. As previously mentioned, future research should directly address the question of hemispheric lateralization using the DAVE stimuli through neuroimaging. Several of the arguments posed could be further supported by information about the affective arousal of the viewers. In particular, pupil dilation and SAM ratings should be measured during the viewing of the DAVE stimuli and findings correlated with hypotheses about the approach and avoidance emotions. Further interpretation of the current findings would benefit from the direct comparison of congruent movies across the two different

allowed for the exploration of gaze behavior during both explicit and implicit judgments of emotion congruence. These results replicate and expand previous knowledge of emotion perception in typically developing individuals, as well as afford researchers and clinicians a foundation from which to explore the social deficits found in a variety of clinical populations.

Results indicated that participants were consistently able to identify angry, fearful, happy, and neutral expressions, as well to determine whether the emotion presented in facial affect matched the emotion presented in vocal prosody. The increased task difficulty of the incongruent conditions was evidenced by the increase in response times, wherein subjects showed a bias toward the visual affective information when asked to determine the emotion portrayed.

Participants consistently showed a preference for the eye region, as opposed to gathering information across both the upper and lower core facial regions, suggesting a feature-based processing approach. Furthermore, during the emotion identification task, the finding that participants were more likely to choose the facially expressed emotion during the incongruent movies is consistent with the increase in fixations to the eye region across all emotion judgment tasks. That being said, further analysis indicated that different patterns exist for each of the emotion conditions. The importance of gleaning information from the eye region appeared to most pronounced for the fearful stimuli, a biologically significant adaptation that is critical to the perception of threat and the source of danger. The inclusion of a neutral emotion condition served as somewhat of a baseline for the current analyses and results showed that participants more often looked at the center (nose) of the stimulus than for other emotion conditions. This provided a nice comparison for the way in which incongruent affective information encouraged the viewer's attention to alternate regions of the face.

participants spent significantly less time eliciting information from any one feature. The preference for the right eye that emerged during the emotion judgment task remained during the match/mismatch task. The automatic perception of fear may have cued participants to fixate on the right eye, a position from which they were able to quickly determine that visual and auditory affective cues were inconsistent (Vuilleumier et al., 2004).

The pattern elicited during neutral movies, unlike the other emotion conditions, was consistent with that predicted for the match/mismatch task. The pattern of fixations was comparable for the congruent and incongruent neutral movies, both in fixation duration and count. As previously stated, the match/mismatch task does not ask participants to label explicitly the emotion perceived. When viewing congruent neutral movies, the participants appear to spend comparable amounts of time on the left eye and the mouth, suggesting that these regions provided equally informative affective cues. The equivalent fixation pattern during the incongruent movies suggests that participants may have found that the lack of affective cues perceived in the visual information registered in a similar manner as during the congruent movies; participants were not cued look for conflicting prosodic cues in other facial regions, but instead relied on information from the auditory modality to confirm that the emotions were incongruent.

Conclusions

The current study provides new insight into the field of emotion perception. Through the direct exploration of gaze behavior, information about the visual scanning approach and visual sampling of facial features was collected during the perception of novel dynamic audio-visual emotion (DAVE) stimuli. These stimuli were digitally altered to allow for the presentation of both congruent and incongruent audio-visual affect. In addition, variation in task instructions

Fixation Patterns: The Effect of Emotion on Match/Mismatch Judgments

As previously discussed, the hypothesis that participants would rely equally on the information from the eyes and mouth due to the overt cuing to integrate facial and vocal affective information was not supported; participants continued to fixate for longer and more often to the eye region. However, the results of exploratory analyses looking at each of the four emotion conditions independently revealed different patterns across emotions that are worth discussing.

Although they were not directly compared, the pattern of fixations that emerged during the viewing of congruent movies appeared to be similar regardless of task instructions, with the exception of the neutral movies. However, the patterns that emerged during the incongruent movies differed according to the task instructions. When presented with incongruent angry and happy movies during the match/mismatch task, the preference for the left eye that was seen in the emotion identification task was absent. When the visual affective information registered as incongruent with the auditory information, occurring early during processing, participants may have been led to look towards other regions in search of evidence to confirm the incongruence of visual and auditory affective cues. These findings could be explored further by comparing fixation patterns for each of the angry- and happy-face combinations (e.g., angry-fearful versus angry-happy); fixation location during the incongruent movies may vary according to the emotion expressed in the auditory modality.

As was the case during the emotion decision task, the results of the match/mismatch task revealed a very different pattern for the fearful movies. Results indicated shorter and fewer fixations during the incongruent fearful compared to congruent fearful movies. This suggests that the overall perception of danger was attenuated by the incongruent auditory affective information; the pattern of fixations remained similar under both movie conditions, yet

potential explanations. First, some participants may still have looked to the right eye because of the automaticity of threat processing, whereas others may have processed the stimulus emotion more globally and relied on information from the left eye. Another interpretation may be that initial fixations were on the right eye because of the automaticity of threat perception, whereas later fixations rested on the left eye as the right hemisphere activated for more global emotion perception (Borod et al., 1998; Vuillemier et al., 2004). Future analyses could examine these explanations by examining the sequence of fixations and by comparing whether participants can be grouped by eye preference.

In contrast to the fearful movies, the pattern of fixations remained similar during the congruent and incongruent angry and happy movies. The increase in fixations during incongruent movies occurred for the regions that were most salient to participants during the congruent movies. As previously noted, the neutral movies were an exception to this statement, albeit to a very small degree. Participants showed an increase in fixation count to the mouth region and decrease to the right eye during the incongruent neutral movies. During the incongruent movies, a neutral facial expression likely provided minimal affective information in terms of facial feature changes and participants may have increased fixations to the mouth in search of more subtle visual speech information, presumed to be consistent with auditory cues. Previous research has suggested that, in the presence of acoustic noise, participant fixations were shorter for both eyes and longer for the nose and mouth (Buchan et al., 2008). As the nose region was not included for analyses comparing the movie conditions, the potential change in fixations between this and other regions cannot be determined. However, as the nose was the primary region of fixation during the congruent movies of the first task, this comparison should be examined in future analyses.

(Kohler et al., 2004). However, fixations to the nose and surrounding regions (e.g., cheeks, wrinkled nose) may have provided additional affective information that was sufficient for accurate emotion identification.

Incongruent Emotion Judgments. As previously discussed, the presence of incongruent vocal affect significantly increased the fixation time devoted to the most salient features of the face. Recall that incongruent movies are referred to by the emotion expressed in the face (e.g., happy incongruent movies refer to those in which the facially expressed emotion is happy and the voice is fearful, angry, or neutral). When examining the fixation pattern for each of the individual incongruent emotion stimuli, there was an overall increase in duration and count of fixations during the incongruent movies for all emotion conditions except neutral. This is consistent with the prediction that the combination of mismatching emotions in the visual and auditory modalities would increase the relative complexity of the stimuli and, therefore, would demand more affective information from facial features.

As was found during analysis of fixation patterns during the congruent movies, the change in fixation pattern between congruent and incongruent movies appears to be most pronounced for the fearful movies. Specifically, the left eye remained the most salient feature during both the congruent and incongruent angry and happy movies. In contrast, participants devoted similar fixation duration and count to each eye during the incongruent fearful movies; this reflects the fact that participants greatly increased their fixations to the left eye during incongruent movies. Rather than increasing fixations to the eye that was previously presumed to be the most salient (right eye), participants increased fixations such that the left eye appeared to provide comparable information during incongruent movies. Perhaps the presence of a conflicting auditory emotion reduced the impact of fear perception, a proposition that has two

In contrast to the neutral movies, the perceptions of congruent angry, fearful, and happy stimuli led to increased fixation duration to the eyes, with fearful stimuli eliciting a pattern much different from the others. During the congruent fearful movies, fixation duration to both eyes and the nose were greater than to the mouth. Similar to the argument posed previously, the perception of danger seemed to encourage fixations to the eye region (most salient), which may actually include the nose in some instances (Green et al., 2003). Previous research indicates that a raised upper eye-lid and raised outer-eyebrows are characteristic and uniquely qualifying features of fearful expressions, respectively (Kohler et al., 2004). Gamer and Buchel (2009) measured eye movements and brain activation during the perception of static fearful faces and found that reflexive gaze movements, originating from a variably located fixation cross, more often changed toward the eye region than the mouth region. This reflexive gaze movement was significantly correlated with amygdala activation for fearful faces, whereas there was no relationship for happy, angry, or neutral emotion conditions. Affective facial cues perceived in the periphery quickly and strongly oriented participants toward the eye region (Gamer & Buchel, 2009).

Unlike the fearful movies, the perception of congruent angry and happy movies relied heavily on affective information from the left eye, with the nose being the second-most preferred region. As a reminder, the left and right eyes are labeled from the perspective of the viewer, such that the left eye refers to the left side of the actor's face. This is in contrast to reports that recognition of happy facial expression depends primarily on the motion of the mouth and its surrounding regions (i.e., upper lip, cheeks; Kohler et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008) and only partially consistent with reports that the recognition of anger is facilitated by furrowed or lowered eyebrows, wrinkled nose, and a depressed lower lip

The fixation patterns to the core facial features also differed significantly across the four types of congruent emotions. Specifically, the difference between congruent neutral and the three other congruent emotion conditions suggests that distinguishing emotionally laden stimuli relied on affective cues from particular facial features. Others have found that judgments of different facial expressions require information from specific facial features (Aviezer et al., 2008; Kohler et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008; Smith, Cottrell, Gosselin, & Shyns, 2005). However, several of these authors presented data gathered during experimental tasks very different from that of the current study, including monitoring eye movements during the identification of static facial expressions presented in context (Aviezer et al., 2008), rating characteristic and unique affective units (AUs) present in static photographs (Kohler et al., 2004), and digitally manipulating facial motion during the emotion identification of video clips ranging from approximately 0-10 seconds long (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008). As such, comparisons with these studies will be made taking measurement differences into account.

The finding that participants primarily fixated on the nose during the neutral movies suggests that they were able to gather sufficient emotional information from that position. Because a neutral facial expression is, in some sense, emotionless, it can be assumed that the minimal changes in facial features detected through peripheral vision was sufficient to signal the absence of an emotional expression. Some have reported a shorter duration and fewer fixations to the feature areas of the face (eyes, nose, mouth) for photographs portraying neutral emotions as compared to angry or fearful emotions, supporting the idea that neutral emotions require fewer fixations to elicit affective information (Green, Williams, & Davidson, 2003).

time may have been devoted to the nose region, such that participants were able to successfully integrate prosodic information (nose, mouth) with visual affective information.

Fixation Patterns: The Effect of Stimulus Affect on Emotion Judgments

Congruent Emotion Judgments. The fixation preference for the core regions of the face, when compared to the periphery, was significantly more pronounced for congruent fearful than both congruent angry and happy stimuli. One might infer that the perception of a potential threat strongly captures the viewer's attention. The importance of eliciting information from the most salient (core) features increased greatly when fear was perceived in another individual, as this may be indicative of the presence of danger. Others have found that participants increased focus upon the features of the face during the perception of threat (Green, Williams, & Davidson, 2003). Those authors, however, reported that the perception of threat arose when viewing both angry and fearful static faces. This difference may be accounted for the emotional arousal elicited in the viewer; as previously discussed, anger is often classified as an approach emotion, whereas fear is classified as an avoid emotion (Davidson et al., 1990). Perhaps viewing dynamic presentations of fearful emotions elicits more threat arousal than does viewing dynamic angry emotions. Using the previously stated argument that fear activated the left hemisphere (i.e., left amygdala) more strongly than angry, this suggests that the fearful movies may have been perceived as more arousing. One approach for further assessing this hypothesis would be to include ratings of arousal for each of the DAVE stimuli. Bradley and Lang (1994) developed a pictorial assessment system called the Self-Assessment Manikin (SAM), which provides one method for gathering participant ratings of pleasure, arousal, and dominance. Through SAM ratings of the affective dimension of arousal, ranging from excited to relaxed, the argument that DAVE stimuli elicited more arousal during fearful than angry movies could be tested.

Applying this theory to the results of the emotion decision task suggests that the transfer of fixations occurred between two regions other than the eyes and mouth. Perhaps the increase to the eye region resulted from a decrease in fixations to the nose or peripheral regions. It may be that participants, when presented with a more difficult emotion identification task, diverted their fixations from facial features that are less affectively salient; for example, the nose, in isolation, undoubtedly provides less affective information than does the mouth. This region may have proved a sufficient vantage point during the congruent movies, as previous research has suggested that the participants may gather necessary affective information from the nose and mouth regions for some emotions (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008).

The evidence from speech perception tasks can be applied to the match/mismatch task in a similar manner. The results indicated that the cue to integrate the face and voice did not cause participants to look more to the mouth, suggesting that the mouth region is not that which is associated with the voice for typically developing individuals, as has been suggested for visual speech information (McGurk & McDonald, 1976). Buchan and colleagues (2007) found that fixations were closer to, but not actually on, the mouth region during speech perception tasks. Moreover, Buchan and colleagues (2007, 2008) explored how the addition of acoustic noise (i.e., multi-talker babble) affected the scanning of faces during a speech detection task. In addition to increasing the duration to the mouth, participants looked longer and more often to the center of the face. Taken together, these findings suggest that the center of the face (i.e., nose) may provide a sufficient point from which participants are able to extract visual speech information. Again, because the fixations to the nose were not included in analysis, this theory cannot be considered directly. However, further analysis may reveal that a sufficient proportion of fixation

they increased fixations to the eye region, resulting in a greater difference between the eye and mouth region, during the incongruent movies. However, this difference resulted primarily from the dramatic increase in the eye region, as there was no decrease from the mouth region during the incongruent movies as compared to the congruent movies. In contrast, fixation patterns were comparable during both the congruent and the incongruent movie conditions when congruence judgments were required. As predicted, participants approached the two conditions similarly, as the decision whether the face matches the voice or not required the same information regardless of movie type. Contrary to the hypothesis about fixation patterns during congruence judgments, however, participants did not fixate for as long or as often on the mouth as the eye region, but rather they still relied heavily on the eyes during the congruence judgments.

These results are similar in the sense that predictions about changes in fixations to the mouth region were not supported. Evidence from speech perception tasks provides a possible interpretation for these findings. Researchers have found that when asked to integrate both auditory and visual speech information and report what they heard, participants did not fixate exclusively on the mouth region, the region frequently believed to be the primary source of visual speech information (Buchan, Pare, & Munhall, 2007, 2008; Pare, Richler, ten Hove, & Munhall, 2003). Instead, participants in these studies also devoted a significant proportion of fixations on the nose. These authors posited that direct foveation of the mouth may not be necessary for the perception of visual speech information; visual speech information may be similarly extracted from the central features of the face and, in some cases, from the eyes (Buchan, Pare, & Munhall, 2008; Pare et al., 2003). It should be noted, however, that because the eye and mouth regions were the only regions compared, fixations to the remainder of the face could not be considered directly in the current study.

Fixation Patterns: Congruent and Incongruent Movies

The results provide a great amount of information about the pattern of fixations, both duration and count, used by individuals during the perception of emotional stimuli. Consistent with previous research, participants fixated significantly longer and more often on the core regions of the face (eyes, nose, mouth) than other regions (Luria & Strauss, 1978; Macworth & Bruner, 1970; Manor et al., 1999; Mertens, Sigmeund, & Grusser, 1993; Stacey et al., 2005). Among these core regions, participants demonstrated a bias toward the eye region when asked to make emotion judgments, in that they looked more often toward the eye region than both the nose and the mouth regions. The Gaze Assumption Hypothesis states that participants rely on the parts of a stimulus that are critical to the task instructions (Lansing & McConkie, 1999). During an emotion decision task, the participants elicited critical information from the eyes, with the nose and mouth regions providing additional information in some cases. The eye region, as predicted, appeared to provide additional affectively salient information during incongruent movies, as fixations to this region increased.

Although they were not compared directly, the visual scanning patterns that emerged during congruent movies appeared to be very similar in both the emotion identification and congruence judgment tasks. As predicted, however, differences in fixation patterns were found when incongruent prosodic information was present; these differences appeared to vary as a function of task instructions. During the emotion judgment task, it was predicted that the incongruent movies would be relatively more complex and less automatic, resulting in an increased reliance on the most affectively salient features of the face. In particular, an increase in fixations was expected for the eyes and a decrease expected for the mouth. The participants did in fact increase their fixation duration and count during the incongruent movies. Specifically,

expressions compared to those with either angry or neutral expressions, whereas there was no difference on fixation to the left side between angry and neutral faces (Bate, Haslam, & Hodgson, 2009).

There are several ways to explore further the lateralization of neural processes underlying emotion perception through eye-tracking measures. One possibility for future research would be to examine the sequence of fixations. If the previous hypothesis is valid, that all emotions are initially processed by the right hemisphere and then the left hemisphere dominates for the perception of fear, then one would expect to find that initial fixations rest on the left side of the stimulus. Subsequent fixations for fearful stimuli would then be focused and remain on the right side of the stimulus. Eye-tracking measure could also be used to evaluate the VH that posits specialized hemispheric roles for approach/avoidance emotions. This theory is based on differences in the mood elicited in the viewer; differences in activation of the anterior frontal cortex suggest the role of affective arousal in the viewer during the appraisal of emotions (Davidson et al., 1990). Measures of gaze behavior provide data regarding the changes in pupil dilation during stimulus perception. Pupil dilation has been used by others as a psychophysiological measure of emotional arousal (Bradley, Miccoli, Escrig, & Lang, 2008). Support for the approach/avoidance theory would come from differential pupil dilation during the avoidance stimuli as compared to the approach stimuli. Differences between angry and fearful emotional expression would support the categorization of these emotions based on arousal as opposed to valence, providing further insight into the differences in lateralization during the current study.

emotions. Specifically, they posited that automatic and global processing of emotions, below the level of consciousness, activates the right amygdala; this suggestion is consistent with the RHH. Following this initial reaction, however, the left amygdala takes over in the processing of emotions on a more cognitive perceptual level. In fact, some have suggested that the rapid processing of fear reaches the amygdala through subcortical projections, bypassing the striate visual cortex (see Vuilleumier & Pourtois, 2007 for a review). Evidence suggests that the amygdala then provides direct feedback to the visual cortex for further processing of emotional faces; researchers found that patients with amygdala lesions did not show increased activation in fusiform gyrus for fearful as compared to neutral faces. In addition, the degree of amygdala sclerosis was inversely related to the differential response to fearful versus neutral faces in ipsilateral visual areas (Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). The findings of the current study may indicate that all emotions were initially processed within the right hemisphere, after which the left amygdala was quickly recruited for a more critical evaluation of negative emotions.

The finding that the angry movies elicited greater fixation time on the left side of the stimulus during approximately one-third of the congruent movies and no preference for either side during the remaining two-thirds of the congruent movies, might explain the inconsistent classification of angry: when evaluating valence effect, angry is sometimes grouped with negative emotions (right hemisphere) and sometimes grouped with approach emotions (left hemisphere). Perhaps angry emotions are processed more equally by both hemispheres than other emotions, with regions recruited for secondary processing depending on the reaction of the viewer (i.e., approach or perception of threat). In fact, some authors have found that viewers spent significantly more time fixating on the left side of the stimulus for novel faces with happy

support for the RHH. However, the greater activation of the left hemisphere during the perception of fearful faces is contrary to this hypothesis.

One group recently conducted a meta-analysis of imaging studies exploring the lateralization of emotion perception. Fusar-Poli and colleagues (2009) reported that the processing of emotions, regardless of valence, activated regions bilaterally in several visuo-limbic areas, including the parahippocampal gyrus and amygdala, fusiform gyrus, lingual gyrus, precuneus, inferior and middle occipital gyrus, posterior cingulate, middle temporal gyrus, and inferior and superior frontal gyri. When comparing emotions on the basis of valence, they found a laterality effect for the negative emotions in the left amygdala more than right amygdala, a finding that is in opposition to the VH. When comparing approach and avoidance emotions, the authors reported a difference in the frontal gyrus; the left inferior frontal gyrus showed greater activation for approach emotions, whereas the right medial and middle frontal gyri showed greater activation for avoidance emotions (Fusar-Poli et al., 2009).

The analyses of Fusar-Poli (2009) provide a potential explanation for the results of this study, particularly the finding that fearful faces elicited preferential fixations of the right side of the stimulus. It is possible that the RHH best explains the overall perception of emotions; the right hemisphere activation explains the greater fixation time on the left side of the stimulus that occurred consistently during happy movies and during a subset of the angry movies. In contrast, the preferential fixation to the right side of the stimulus during the fearful movies is commensurate with the finding that negative emotions elicited greater activation in the left amygdala more than the right (Fusar-Poli et al., 2009). This argument requires consideration of the role of the amygdala in the perception of potential sources of threat. Fusar-Poli et al. (2009) suggested that the left and right amygdala might play different roles in the processing of

were significantly impaired on emotion identification tasks, using static photographs and audio-taped sentences in eight positive and negative emotions when compared to patients with left-hemisphere brain-damage. The RHH is further supported by findings that the neural network of face processing is predominantly right-sided (Kanwisher, McDermott, & Chun, 1997).

The second hypothesis of emotion perception is the Valence Hypothesis (VH), which proposes that both hemispheres have a role in emotion perception, but that each is specialized for specific emotions. This hypothesis posits that positive emotions (e.g., happy) are processed by the left hemisphere and that negative emotions (e.g., fearful, angry) are processed by the right hemisphere; it has been found that right-hemisphere damage reduces the usually superior discrimination of negative emotions (e.g., sad) and that left-hemisphere damage reduces the usually superior processing of positive emotions (e.g., happy; Adolphs, Jansari, & Tranel, 2001). The valence hypothesis has also been differentiated on another dimension, on that has to do more with the mood elicited in, or the affective arousal of, the viewer. This approach/withdrawal theory proposes that approach emotions (e.g., angry, happy) are processed in the left hemisphere and the withdrawal emotions (e.g., fearful, sad) are processed in the right hemisphere (Davidson, Saron, Senulis, Ekman, & Friesen, 1990).

The findings of the current study, that happy and angry movies resulted in longer fixation time to the left side of the stimulus (i.e., appear to activate the right hemisphere more than the left) and that fearful resulted in longer fixation time to the right side of the stimulus (i.e., appears to activate the left hemisphere more than the right), do not map clearly onto either the RHH or the VH. The greater involvement of the right hemisphere in the processing of happy and angry movies, as inferred from the preferential fixation on the left side of the stimulus, provides some

relationship between directional eye movements and hemispheric activation of emotion perception often employ subjective methods for recording eye movements, such as when an experimenter visually observes and manually records the number of eye movements made in each direction (e.g., Asthana & Mandal, 2001; Graves & Natale, 1979). In a study attempting to resolve contradictory findings in the conjugate lateral eye movement literature, Gur (1975) conducted a within-subjects study to examine different methodologies. Gur (1975) reported that only when the experimenter was seated behind the participant and eye movements were recorded by video was the direction of conjugate lateral eye movements dependent on type of question presented: verbal items (i.e., left hemisphere activation) caused right-sided movements, whereas spatial items (i.e., right hemisphere activation) caused left-sided movements.

The following discussion assumes that the direction of conjugate lateral eye movements indicates activation of the contralateral cerebral hemisphere, given that similar to Gur (1975), eye movement recordings are made objectively (also see Kinsbourne, 1972). Findings from an objective measure of lateral eye movements (i.e., visual field bias) are discussed in the context of theories about the lateralization of emotion perception, as illustrated by direct measures of hemispheric lateralization (i.e., fMRI studies). However, in light of the speculative nature of such a connection, the current discussion would benefit from future corroboration through neuroimaging in which visual field bias can be directly related to hemispheric activation.

There are two primary theories about the lateralization of emotion perception and potential effects of valence or affective interpretation. The first is the Right-Hemisphere Hypothesis (RHH), which states that right hemisphere is specialized for the processing and recognition of all emotions (Borod et al., 1998; Tamietto, Corazzini, de Gelder, & Geminiani, 2006). Borod and colleagues (1998) found that patients with right-hemisphere brain-damage

differentiate emotionally salient information. In contrast, dynamic changes perceived through peripheral vision may signal changes in configuration that are unique to a particular emotion, thus eliminating the need to scan the face for such information (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008). One method for testing this hypothesis would be to examine the processing approach of the DAVE stimuli under static conditions, with the expectation that the elimination of configural changes would result in increased visual scanning of facial regions.

Review of the processing approach data showed that twice as many participants demonstrated a configural processing approach ($n=12$) during the incongruent movies of the emotion identification task as during the congruent movies ($n=6$; see Table 9). Although this remains a significant minority of participants, this pattern is opposite of that predicted and deserves further attention. The reduced automaticity of the incongruent emotion judgments may have led participants to scan to other facial regions in search of cues that would resolve the incongruence. One future direction for research could analyze whether this subset of participants were also more likely to respond with the vocally portrayed emotion during incongruent movies.

Lateralization of Emotion Perception

The results of the current study indicated that participants fixated for longer periods of time on the left side of the stimuli significantly more than right during the angry and happy congruent movies, as well as a portion of the happy incongruent movies. Several researchers have theorized the role of the right and left hemisphere in various aspects of emotion perception and researchers often use asymmetric viewing strategies or visual field advantage as evidence of differential hemispheric involvement. It should be noted, however, that the degree to which we are able to infer differences in hemispheric activation from the directional biases of gaze fixations remains questionable. For example, researchers drawing conclusions about the

meaning that rather than only perceiving face regions within their foveal vision, participants were able to process facial information present in their peripheral vision.

A second possible cause for the contradictory findings related to processing approach may be related to the requirements of the tasks in the current study. Several of the previous studies that highlight the importance of configural information are studies of face recognition, whereas the current study addresses processing strategies in facial emotion recognition. In order to recognize a person from one time point to another, the use of specific information (second-order relational information) is critical to distinguishing one face from another. On the other hand, the recognition of facial emotions is believed to be more universal. Certain facial feature changes are consistent in the presentation of specific emotions; looking to those features (the eyes, the mouth) and perceiving specific orientations of the features may be clue enough to the emotion presented.

Processing approaches may also differ according to whether stimuli are static or dynamic. In a static stimulus, and often under longer viewing times, participants may be required to scan the face more in order to process configural information and the shape of individual features. In contrast, a dynamic stimulus inherently provides configural information that changes in a temporal fashion. As discussed earlier, the dynamic presentation of facial expressions improves recognition; participants benefit from the perception of these changes over time, a skill that typically developing adults have well established. It is possible that participants perceive changes in facial configuration through fixations on one particular feature. For example, two emotions for which identification relies on the shape of the eyes can generally be distinguished on the basis of unique eye configurations (e.g., angry, fearful; Kohler et al., 2004). In a static stimulus, these configurations are stable across time and participants may need to scan this or other regions to

processed the special configuration of the facial features. However, it appears that the participants in this study relied more heavily on the eye region, meaning that participants used a featural approach rather than a configural approach.

Despite the contradiction between the current results and those of previous studies, there are several points to consider when reconciling the differences, namely the type of data collected, the task instructions, and the type of stimuli used. With regard to the data collected, previous researchers have used different perceptual paradigms for assessing configural and featural processing than that which was used in the current study. The studies previously cited have looked at recognition performance (facial and emotion) as it relates to the inversion effect, composite effect, and part-whole stimuli; participants consistently showed a recognition advantage for upright faces, when viewing properly aligned facial features, and when viewing features in the context of the whole face. Using these paradigms, researchers concluded that the improved recognition was due to the availability of configural information. Given these differences in measurement, the current findings may not fully contradict these previous findings; it may be that the participants in the current study gathered configural information that could not be measured by examining the fixation points provided with eye-tracking measures.

Some research has indicated that participants process all of the features of the face in parallel, as opposed to representing these features individually (Bradshaw & Wallace, 1971; Farah et al., 1998). If faces are represented in a holistic or gestalt form during processing, making inferences about the processing approach from individual fixations may be misleading. To be discussed in more detail in the following sections, there was evidence to suggest that participants perceived information from parts of the face on which they did not directly fixate,

al., 2008; Hecht & Reiner, 2008; Koppen, Alsius, & Spence, 2008) and visual affective information has a greater influence on prosodic information than does the reverse influence (de Gelder & Vroomen, 2000; Massaro, 1998; Massaro & Egan, 1996).

Processing Approach

Numerous studies have been conducted to examine the process by which individuals elicit information from human faces. Strategies used during face identification, recognition, and categorization have been identified and several researchers have stressed the importance of configural information in accurate performance (Diamond & Carey, 1986; Rhodes, 1988; Rhodes, Brake, & Atkinson, 1993; Valentine, 1988). Configural information has also been implicated in the accurate perception of facial expressions (Gross, 2005; Pelphrey et al., 2002; White, 1999). As a result, it was hypothesized that participants in the current study would show evidence of a configural processing approach during the perception of congruent emotion stimuli. In addition, it was hypothesized that the overt cue to integrate visual and auditory affective information during the match/mismatch task would lead participants to use a configural strategy when making congruence decisions. In contrast, it was predicted that the presentation of the incongruent movies during an emotion decision task would cause participants to increase their reliance on the most salient visual affective cues, which would lead them to use a featural processing strategy in order to focus on the eyes during these trials. Results indicated that participants predominantly used a featural processing approach during all congruent and incongruent movies, during both the emotion decision and match/mismatch tasks. In the current study, a configural processing approach was classified using fixations to both the eye region (both left and right eyes combined) and the mouth region. This would have indicated that participants scanned the whole of the face, from the upper half to the lower half, and presumably

gyri and right inferior frontal gyrus has been associated with decoding of emotional prosodic cues (Mitchell, 2003), whereas the left inferior frontal gyrus has been associated with the syntactic processing in sentence comprehension (Bilenko, Grindrod, Myers, & Blumstein, 2008). Taken together, these findings suggest that the perception of incongruent bimodal information leads to additional processing in anatomical regions associated with unimodal perception, particularly for the non-dominant modality (speech comprehension). As applied to the current study, Robins, Hunyadi, and Schultz (2009) found increased activation in the posterior superior temporal sulcus during the perception of congruent DAVE stimuli when compared to unimodal presentation of the DAVE stimuli. Perhaps the perception of incongruent stimuli would lead to further processing in regions associated with decoding facial expressions (e.g., fusiform gyrus; Gauthier et al., 1998; Gauthier & Tarr 1997; Kanwisher, McDermott, & Chun, 1997; Kanwisher & Moscovitch, 2000) and, more importantly, the decoding of prosody (lateral temporal lobes; Mitchell, 2003).

The question remains as to how participants arrived upon an answer when they are faced with incongruent information. In real-life social situations, individuals have the opportunity to blend visual and auditory speech information or affective cues, in a manner illustrated by the McGurk Illusion (McGurk & MacDonald, 1976). However, because of the forced-choice format of the current tasks, participants were unable to provide a blended emotion response. When the facial expression and the vocal prosody were mismatching and participants were asked to label the emotion presented, participants in the current study were more influenced by the visual modality. In other words, participants consistently responded with judgments matching the facial expression, when compared to those matching the vocal prosody. These findings replicate those that have found that visual information is dominant during visual-auditory stimuli (Collignon et

Despite the increased difficulty of these decisions, participants were able to identify correctly whether the emotion presented in the face matched the emotion in the voice. This finding held true for both the congruent and incongruent movies. Contrary to the prediction that response times during these match/mismatch judgments would be approximately equal during both congruent and incongruent movies, participants in the current study were slower to respond to the incongruent movies. This provides further evidence for the relative automaticity of basic emotion integration and recognition; congruent visual and auditory information was quickly integrated and correctly identified, whereas the relatively more complex task of integrating and evaluating incongruent visual and auditory information required a greater amount of time. Previous research has shown that the integration of bimodal information occurs early in processing (Hietanen et al., 2004; Pourtois et al., 2000). As a whole, these results suggest that when participants perceive incongruence at the early stage of integration, both the identification of emotion and match/mismatch decisions are delayed as a function of the extraordinary nature of the stimuli.

Perhaps when participants are faced with bimodal information that does not quickly register as belonging to a particular affective category, they recruit additional cognitive resources to arrive upon an affective judgment. For example, the perception of congruent prosody and lexico-semantic information, relative to prosody-only, has shown bilateral superior temporal gyri and left middle temporal gyrus, as well as the left ventrolateral prefrontal cortex (Mitchell, 2006). However, when participants were required to identify emotional prosody in the presence of incongruent lexico-semantic information, there was increased activation in the left inferior frontal gyrus, left ventrolateral prefrontal cortex, bilateral middle temporal gyri, right superior temporal gyrus, as well as other regions. Activation in the right middle and superior temporal

emotion perception deficits that affect a number of clinical populations, including ASD, social anxiety, and schizophrenia.

Behavioral Responses and Response Time

Social interactions require the integration of both visual and auditory information, a process that occurs quickly and, seemingly automatically, during the perceptual stage of processing (Hietanen et al., 2004; Pourtois et al., 2000). When presented with congruent stimuli and required to label explicitly the emotion, participants were accurate in their responses. This is consistent with previous research showing that typically developing individuals are consistently able to identify basic emotions that are recognized across cultures (e.g., Ekman & Friesen, 1975). In particular, researchers have posited that dynamic emotion stimuli, as compared to static stimuli, improve the perception of facial expressions by providing critical temporal information, particularly in the case of subtle emotional expressions (Ambadar, Schooler, & Cohn, 2005; Hess & Kleck, 1995).

We sometimes encounter situations in which we receive social cues that are seemingly inconsistent with one another. We then have to decipher these perplexing cues and make a decision about how to proceed in the social interaction. A portion of the current stimuli allowed for the simulation of this relatively more complex emotion perception, as some stimuli were digitally altered to contain mismatching affective cues in facial expression and vocal prosody. The finding that participants were slower to respond to the incongruent movies than congruent movies in an emotion decision task provides evidence of the increased difficulty and decreased automaticity of such responses. These findings are commensurate with other reports of increased response times during incongruent emotion tasks (Hietanen, Leppanen, Illi, & Surakka, 2004; Massaro & Egan, 1996; Pell, 2005).

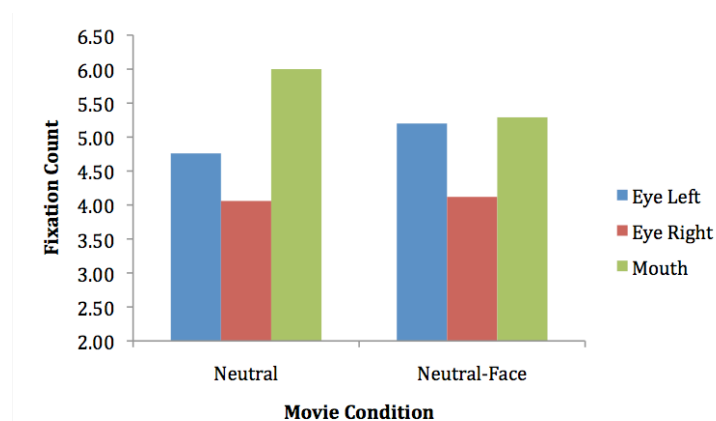


Figure 26. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 3.

Discussion

This study provides information about the pattern of fixations used by typically developing individuals during emotion perception. Among the core facial regions, participants elicited critical information from the eyes, with the nose and mouth regions providing additional information in some cases. The core regions of the face, as predicted, appeared to provide affectively salient information during incongruent movies, as fixation time to both the eye and mouth regions increased compared to fixation times during congruent movies. Measurements of gaze behavior during the perception of emotion stimuli are one of the most direct ways of determining the manner in which individuals approach these emotion perception tasks. The current study examined the gaze behavior of a group of typically developing individuals when viewing emotional movie clips. In contrast to the static photographs, this study used novel dynamic emotion stimuli that allowed for exploration of ways in which bimodal emotional information affected perception. The use of these digitally altered stimuli further allowed for the presentation of facial expression and prosody that were sometimes affectively incongruent. This relatively more difficult bimodal combination provided information, in the form of changed fixation patterns and processing approaches, that may increase our understanding about the

Neutral Movies. The final 2x3 repeated-measures ANOVA comparing fixation pattern was between congruent neutral and incongruent neutral-face movies (Table 20). Analyses revealed no significant main (Movie, Region) or interaction (Movie x Region) effects for either *fixation duration* ($ps=.89$, $.58$, and $.32$, respectively) or *fixation count* ($ps=.72$, $.23$, and $.13$, respectively; Figures 26 and 27). These results provide support for the hypothesis that participants rely on the eye and mouth regions similarly during the congruent and incongruent neutral movies.

Table 21

Descriptive statistics of fixation duration and count during neutral movies of Task 3

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Neutral	2.82 (3.39)	2.09 (2.26)	2.73 (2.74)	4.76 (4.90)	4.06 (4.12)	6.00 (4.60)
Neutral-Face	2.82 (3.04)	2.33 (2.68)	2.53 (2.49)	5.20 (4.56)	4.12 (4.14)	5.29 (4.38)

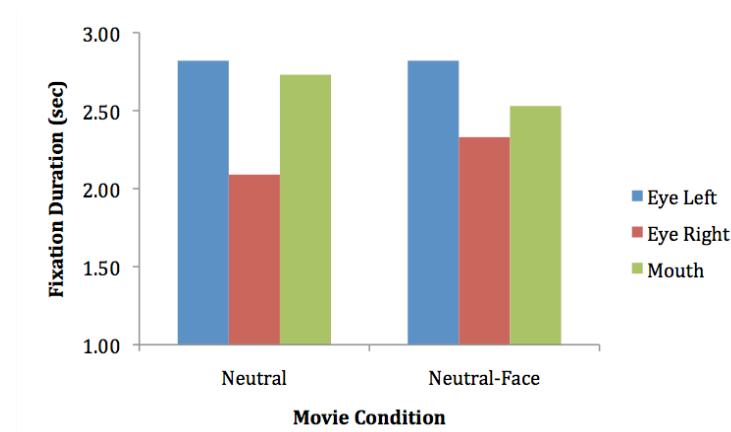


Figure 25. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 3.

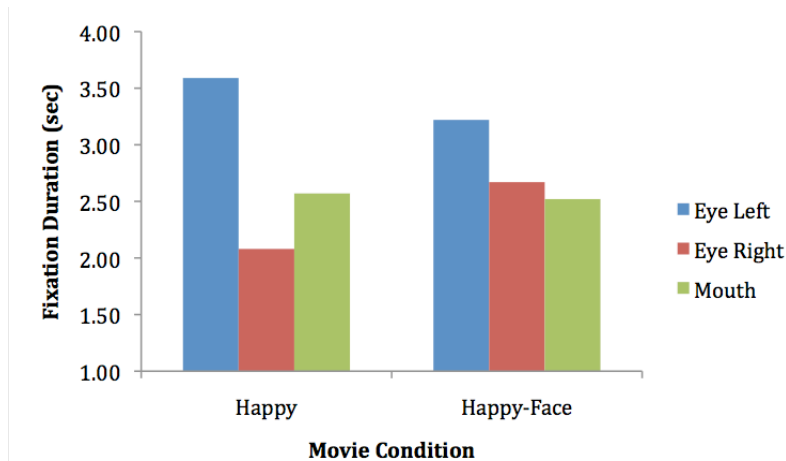


Figure 23. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 3.

Analyses comparing *total fixation count* during happy and happy-face movies, revealed a significant Movie x Region interaction, $F(2,100) = 5.17$, $p = .01$, $\eta_p^2 = .09$. There was no main effect of Movie ($p = .70$) or Region ($p = .23$). Contrasts indicated an interaction when comparing congruent to incongruent movies for the right eye and mouth, $F(1,50) = 8.01$, $p = .01$, $\eta_p^2 = .14$. Figure 25 shows that this difference resulted from the significant increase in fixations to the right eye during the incongruent movies. These findings only partially support the hypothesis because participants rely more equally on the eye and mouth regions during the incongruent movies.

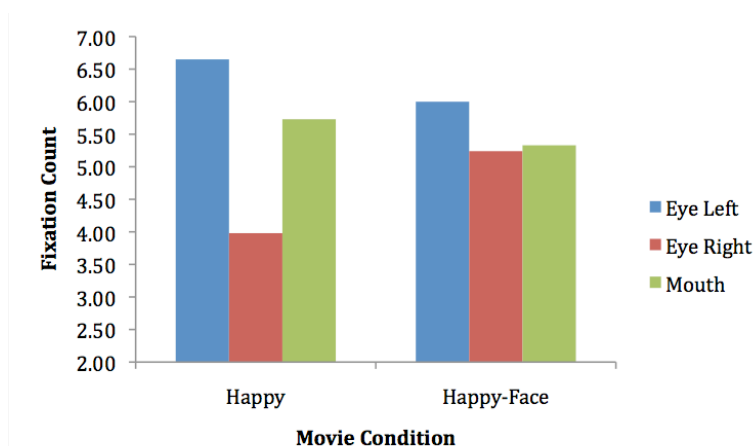


Figure 24. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 3.

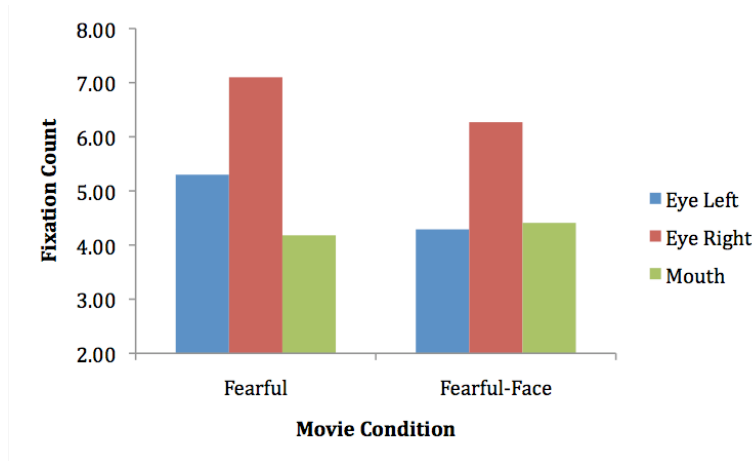


Figure 22. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 3.

Happy Movies. When comparing *total fixation duration* during congruent happy and incongruent happy-face movies, the interaction effect of Movie x Region reached significance, $F(2,100)= 3.86, p= .02, \eta_p^2= .07$ (Table 19). Main effects of Movie and Region failed to reach significance ($p=.58$ and $p=.24$, respectively). Contrasts show that there is an interaction between congruent and incongruent movies for the right eye compared to the mouth, $F(1,50)= 4.54, p= .04, \eta_p^2= .08$. This interaction (Figure 24) reflects a significant increase in fixation duration for the right eye during the incongruent movies. The hypothesis was partially supported by the decreased difference between the eye and mouth regions during the incongruent movies.

Table 20

Descriptive statistics of fixation duration (sec) and count during happy movies of Task 3

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Happy	3.59 (3.48)	2.08 (2.29)	2.57 (2.68)	6.64 (5.01)	3.98 (4.15)	5.73 (5.64)
Happy-Face	3.22 (3.64)	2.67 (2.73)	2.52 (2.81)	6.00 (5.75)	5.24 (5.31)	5.33 (4.86)

for the hypothesis, in that the fixation duration to the eye and mouth regions are more equal during the incongruent movies than the congruent movies.

Table 19

Descriptive statistics of fixation duration (sec) and count during fearful movies of Task 3

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration			Count		
Fearful	2.75 (2.79)	4.32 (3.31)	1.70 (1.96)	5.33 (4.27)	7.10 (5.40)	4.18 (4.65)
Fearful-Face	2.42 (2.73)	3.42 (3.20)	2.03 (2.23)	4.29 (4.17)	6.27 (5.46)	4.41 (4.16)

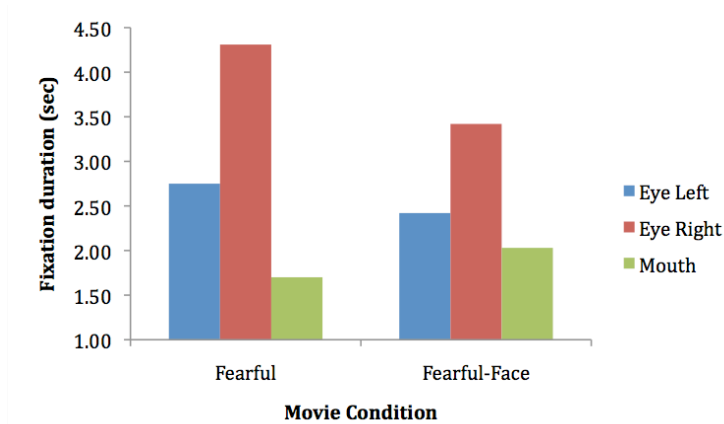


Figure 21. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 3.

Results comparing *total fixation count* during fearful and fearful-face movies indicated a main effect of Movie, $F(1,50)= 5.70, p=.02, \eta_p^2= .10$, with a significantly greater number of fixations during congruent than incongruent movies. There was also a main effect of Region, $F(1.74,86.86)= 3.66, p= .04, \eta_p^2= .07$, and contrasts indicated that the right eye was fixated on a significantly greater number of times than the mouth, $F(1,50)= 5.16, p= .03, \eta_p^2= .09$. The absence of an interaction effect ($p=.20$) fails to provide support for the hypothesis (Figure 23).

incongruent than congruent movies. This change reflects a significant decrease in fixations to the left eye and an increase in fixations to the mouth, such that the mouth was the region fixated on most frequently in the incongruent angry-face movies. This fails to support the hypothesis in that the pattern of fixations differs between congruent and incongruent movies.

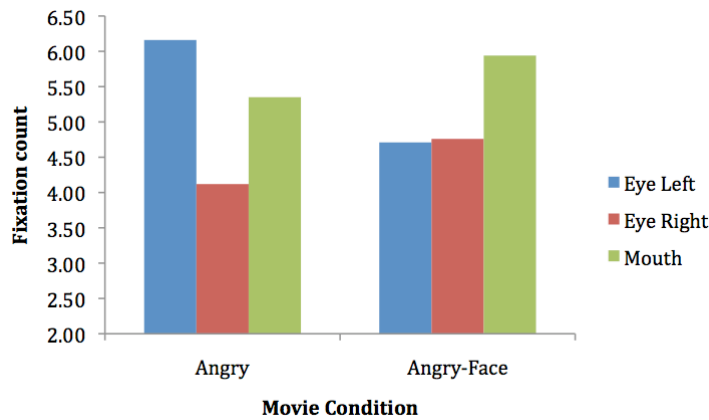


Figure 20. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 3.

Fearful Movies. When comparing the *total fixation duration* during the congruent fearful and incongruent fearful-face movies, a main effect of Movie, $F(1,50)=6.78, p=.01, \eta_p^2=.12$, indicated greater fixation duration during the congruent than incongruent movies (Table 18). There was also a main effect of Region, $F(2,100)=6.30, p<.001, \eta_p^2=.11$, and contrasts indicated that fixation duration to the right eye is significantly greater than to the mouth, $F(1,50)=10.16, p=.002, \eta_p^2=.17$. A significant Movie x Region interaction, $F(2,100)=5.22, p=.01, \eta_p^2=.10$, denotes that fixation duration to the predefined facial regions varies by movie condition. Contrasts revealed interactions between congruent and incongruent movies for the right eye compared to the mouth, $F(1,50)=10.74, p=.002, \eta_p^2=.18$. Looking at the interaction graph in Figure 22, the difference in fixation duration between the right eye and mouth appears to be significantly less during the incongruent than the congruent movies. This fails to provide support

comparing congruent to incongruent for the left eye compared to the mouth, $F(1,50)= 4.50$, $p=.04$, $\eta_p^2= .08$. The interaction graph (Figure 20) shows that the difference between fixation duration to the left eye and mouth is significantly less during the incongruent movies than during the congruent movies. These results fail to support the hypothesis: fixation durations to the eye and mouth regions were more similar during the incongruent movies than the congruent movies.

Table 18

Descriptive statistics of fixation duration (sec) and count during angry movies of Task 3

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Angry	3.47 (3.78)	2.42 (2.64)	2.67 (2.87)	6.16 (5.29)	4.11 (4.24)	5.35 (5.37)
Angry-Face	2.80 (2.96)	2.80 (2.72)	2.79 (2.94)	4.71 (3.85)	4.76 (4.35)	5.94 (5.49)

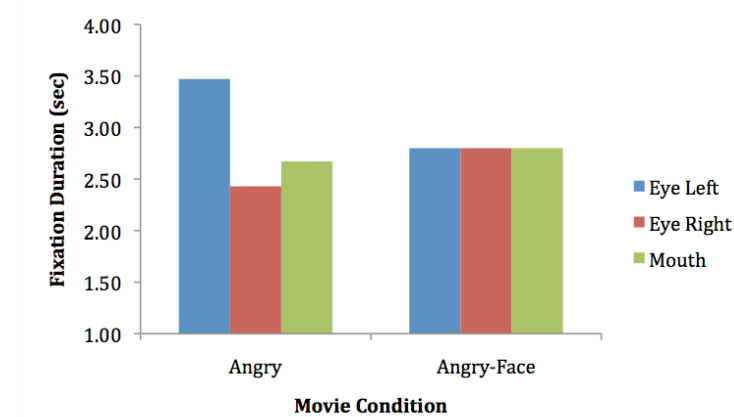


Figure 19. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 3.

Comparison of *total fixation count* during angry and angry-face movies revealed a significant interaction of Movie x Region $F_{(2,100)}= 8.26$, $p<.001$, $\eta_p^2= .14$. Main effects of Movie and Region were again non-significant ($p=.75$ and $p=.44$, respectively). Contrasts revealed a significant interaction between congruent and incongruent movies for left eye compared to mouth, $F_{(1,50)}= 9.13$, $p=.004$, $\eta_p^2= .15$. The interaction graph (Figure 21) shows that the difference in number of fixations between the left eye and mouth is significantly greater for the

regions would remain consistent during both the congruent and incongruent movies partially supported the hypothesis. These findings were explored further in the subsequent analyses examining each emotion condition separately.

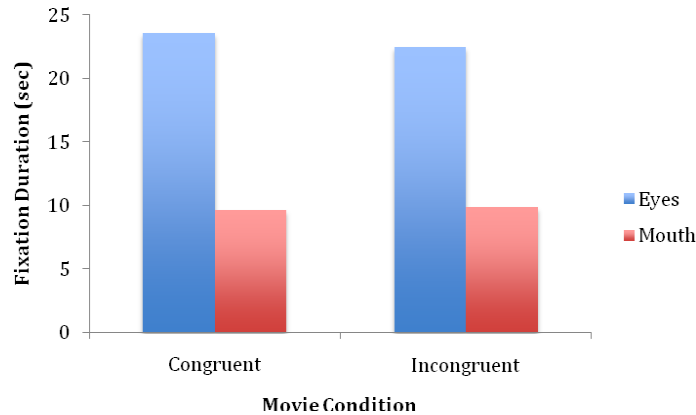


Figure 17. Total fixation duration to the two areas of interest during congruent and incongruent movies of Task 3.

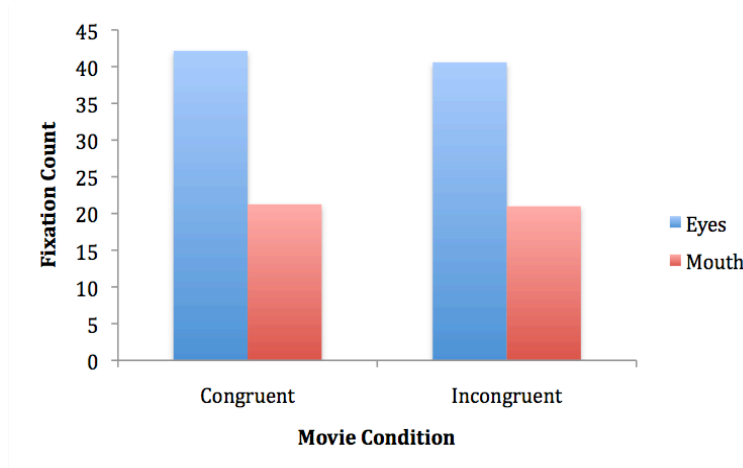


Figure 18. Total fixation count to the two areas of interest during congruent and incongruent movies of Task 3.

Angry Movies. In the analysis comparing the *total fixation duration* during congruent angry and incongruent angry-face movies, analyses revealed a significant Movie x Region interaction, $F(2,100)= 4.47$, $p= .01$, $\eta_p^2= .08$ (Table 17). The main effects of Movie and Region were not significant ($p=.50$ and $p=.70$, respectively). Contrasts indicated interactions when

equal to the *total fixation duration* and *fixation count* devoted to the mouth region, and that this pattern would remain constant during the viewing of both congruent and incongruent movies. The fixation patterns for the two movie conditions were compared directly using the congruent movies of Task 3 to the incongruent movies of Task 3 in a series of 2x3 and 2x4 (Movie x Region) repeated-measures ANOVAs. In the same manner as done for Hypothesis 6, the ANOVAs for the current hypothesis were initially done for all emotions and both eyes combined and then for each emotion individually, comparing left and right eye separately.

Analysis of *fixation duration* between the congruent and incongruent movies of Task 3 revealed a significant main effect of Region, $F(1,50)=14.94, p<.001, \eta_p^2=.23$ (Table 16). Participants spent a significantly greater amount of time fixating on the eye region as compared to the mouth region (Figure 18). The main effect of Movie and the Movie x Region interaction were not significant ($p=.19$ and $p=.09$, respectively).

Table 17

Descriptive statistics of fixation duration (sec) and fixation count during congruent and incongruent movies of Task 3

Condition	Region			
	Eyes	Mouth	Eyes	Mouth
	Duration (sec)		Count	
Congruent	23.54 (17.27)	9.67 (9.46)	42.16 (29.69)	21.25 (18.41)
Incongruent	22.48 (16.46)	9.87 (9.79)	40.59 (29.03)	20.98 (17.58)

Similarly, analysis of *fixation count* revealed a significant main effect of Region, $F(1,50)=12.40, p=.001, \eta_p^2=.20$ (Figure 19). Results indicated greater fixations on the eye region as compared to the mouth region. The main effect of Movie and the Movie x Region interaction failed to reach significance ($p=.19$ and $p=.41$, respectively). The hypothesis that participants would spend approximately the same amount time and number of fixations on the eye and mouth regions was not supported. However, the finding that the difference between the eye and mouth

and voice and, therefore, would increase scanning of both the eye and mouth region. As a result, it was predicted that more participants would utilize a configural facial processing strategy than featural strategy during the viewing of both congruent and incongruent movies. The congruent movies of Task 3 were directly compared to the incongruent movies of Task 3. The number of participants exhibiting configural and featural facial processing strategies for each of the experimental conditions can be found in Table 15.

Table 16

Proportion of participants utilizing configural or featural processing strategies of Task 3

Condition	Processing Strategy		<i>p</i>
	Configural	Featural	
Task 3 Congruent All Emotions	4	47	<.001
Task 3 Angry	3	48	<.001
Task 3 Fearful	6	45	<.001
Task 3 Happy	5	46	<.001
Task 3 Neutral	2	49	<.001
Task 3 Incongruent All Emotions	5	45	<.001
Task 3 Angry Face	7	44	<.001
Task 3 Fearful Face	6	45	<.001
Task 3 Happy Face	6	45	<.001
Task 3 Neutral Face	3	48	<.001

Sign tests show that a greater proportion of participants utilized a featural facial processing strategy than a configural strategy during all congruent and incongruent movie conditions, with combined and separated emotions. These findings fail to support the hypothesis that participants would be pulled to rely on more features of the face when making match/mismatch judgments.

Hypothesis 9 – Task 3 Fixation Pattern

It was hypothesized that, because participants would be overtly cued to integrate affective information from the face (i.e., eyes) and voice (i.e., mouth) during match/mismatch judgments, the *total fixation duration* and *fixation count* devoted to the eye region would be approximately

Contrasts revealed interactions when comparing congruent and incongruent movies for the right eye compared to the mouth, $F(1,50)=6.54, p=.01, \eta_p^2=.12$. The difference between the right eye and the mouth is significantly less in the incongruent movies when compared to the congruent movies. The hypothesis was not supported in the neutral movie condition.

Hypothesis 7 – Task 3 Behavioral Responses and Reaction Time

In order to test the hypothesis that participants, when viewing both congruent and incongruent movies, would be able to identify correctly whether the emotion portrayed in the actor's face matches the emotion portrayed in the actor's voice, the percentage of correct responses was calculated across all movies of Task 3 ($n=64$). Results indicated that participants responded correctly to 81.42% of the movies. It was also predicted that participant response times would be approximately equal when deciding whether the facial emotion matched the vocal emotion during both congruent and incongruent movies. Results of a paired samples t-test failed to support this hypothesis and indicated that incongruent movie response times ($M=760.82$ ms, $SD=263.03$ ms) were significantly longer than congruent movie response times ($M=589.29$ ms, $SD=189.36$ ms), $t(49)=-7.75, p<.001$. For the previous analysis, reaction time data was unavailable for one participant.

Hypothesis 8 – Task 3 Processing Approach

To examine the nature of the participants' scanning patterns, differences in facial processing strategies were assessed in the same manner as done for previous hypotheses. Separate sign tests for all congruent movies, all incongruent movies, four separate congruent emotion conditions, and the four separate incongruent emotion conditions, were conducted in order to compare the number of configural processors to the number of featural processors. The current hypothesis stated that participants would be cued to integrate information from the face

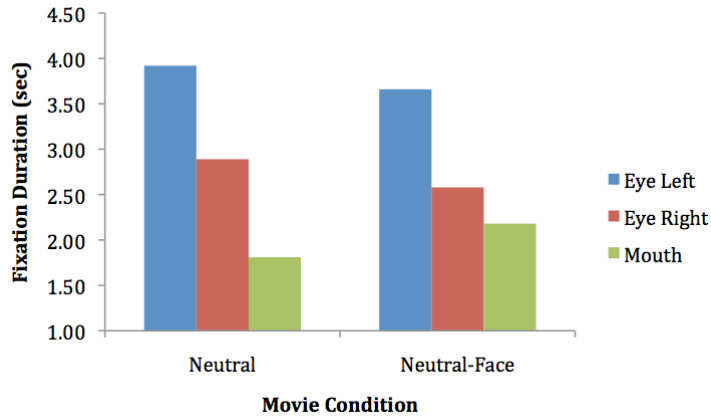


Figure 15. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 2.

Comparison of *total fixation count* during congruent neutral and incongruent neutral-face movies indicated a significant main effect of Region, $F(1.64, 82.02) = 54.79$, $p = .05$, $\eta_p^2 = .06$. Contrasts revealed that a significantly greater number of fixations were devoted to the left eye when compared to the mouth, $F(1, 50) = 4.73$, $p = .03$, $\eta_p^2 = .09$. The main effect of Movie was non-significant ($p = .75$). There was also a significant Movie x Region interaction, $F(1.76, 88.09) = 15.77$, $p = .03$, $\eta_p^2 = .07$ (Figure 17).

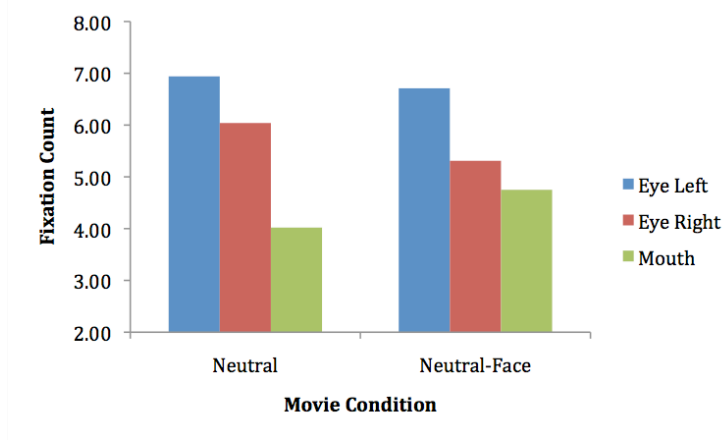


Figure 16. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 2.

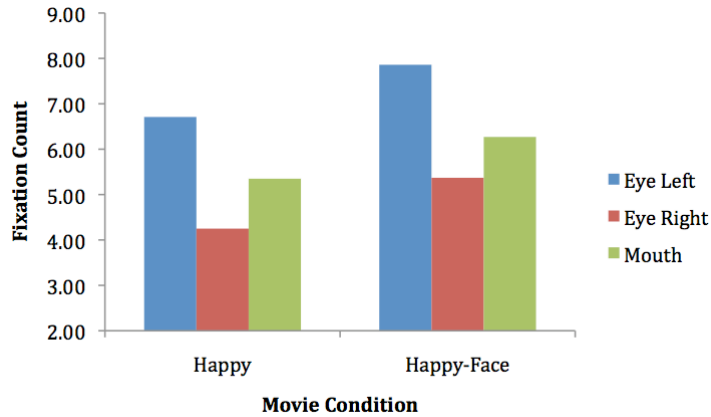


Figure 14. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 2.

Neutral Movies. The final analysis compared the congruent neutral and incongruent neutral-face movies. Analysis of *total fixation duration* revealed a significant main effect of Region, $F(1.76, 88.19) = 4.01, p = .03, \eta_p^2 = .07$ (Table 14). The main effect of Movie failed to reach significance ($p = .48$). Contrasts indicated that total fixation duration was significantly greater for the left eye than the mouth, $F(1, 50) = 9.58, p = .02, \eta_p^2 = .11$. The absence of an interaction effect ($p = .06$) for the neutral and neutral-face movie conditions fail to support the hypothesis that incongruent movies would result in differential changes to the eye and mouth regions (Figure 16). The hypothesis was not supported for the neutral movie condition, as the overall fixation duration for both the eye and mouth regions remained similar during the congruent and incongruent movies.

Table 15

Descriptive statistics of fixation duration and count during neutral movies of Task 2

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Neutral	3.92 (3.76)	2.89 (2.45)	1.81 (2.13)	6.94 (5.66)	6.04 (5.00)	4.02 (3.52)
Neutral-Face	3.66 (3.87)	2.58 (2.70)	2.18 (2.57)	6.71 (5.92)	5.31 (4.99)	4.75 (4.08)

Happy Movies. With regard to congruent happy and incongruent happy-face movies, analyses revealed significant main effects of Movie for both *fixation duration*, $F(1,50)=21.06$, $p<.001$, $\eta_p^2=.30$, and *fixation count*, $F(1,50)=18.82$, $p<.001$, $\eta_p^2=.27$, with greater fixation duration and number of fixations during the incongruent movies than during the congruent movies (Table 13, Figures 14 and 15). Main effects of Region (*duration* $p=.07$; *count* $p=.06$) and Movie x Region (*duration* $p=.55$; *count* $p=.90$) interactions did not reach significance, although the main effect of region approached significance. The results indicating an overall increase in length and frequency of fixations during the incongruent movies provide partial support for the hypothesis.

Table 14

Descriptive statistics of fixation duration (sec) and count during happy movies of Task 2

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Happy	3.52 (3.72)	2.15 (2.30)	2.59 (2.63)	6.71 (5.39)	4.15 (4.48)	5.35 (4.40)
Happy-Face	4.13 (3.98)	2.52 (2.11)	2.91 (2.89)	7.86 (5.80)	5.37 (4.26)	6.27 (4.91)

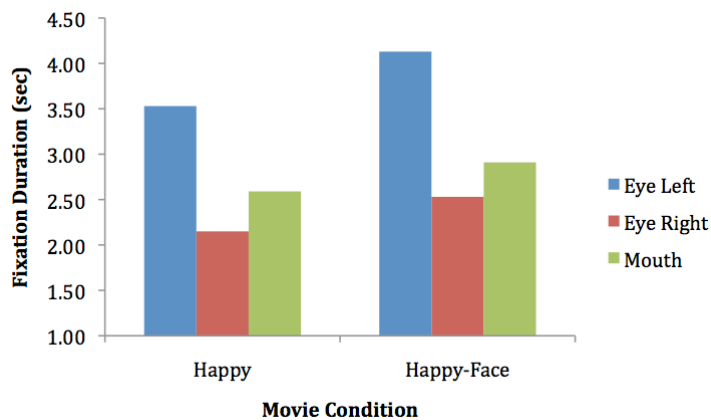


Figure 13. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 2.

greater during incongruent movies than congruent movies. These findings provide support for the hypothesis when considering the fearful movie condition, in that fixation duration increased for the left eye and decreased for the mouth

Comparison of *total fixation count* for fearful versus fearful-face movies indicated significant main effects of Movie, $F(1,50)= 7.86, p= .01, \eta_p^2= .14$, and of Region, $F(2,100)= 5.84, p=.004, \eta_p^2= .11$. Contrasts showed significantly more fixations during incongruent movies when compared to congruent movies and for both the left eye, $F(1,50)= 4.67, p= .04, \eta_p^2= .09$, and right eye, $F(1,50)= 10.93, p=.002, \eta_p^2= .18$, when compared to the mouth. There was also a significant Movie x Region interaction, $F(2,100)= 10.73, p< .001, \eta_p^2= .18$. Contrasts revealed a significant interaction when comparing congruent and incongruent movies for the left eye compared to the mouth, $F(1,50)= 16.14, p< .001, \eta_p^2= .24$. As seen in Figure 13, the difference in fixation count to the left eye and the mouth is significantly less during the congruent movies when compared to the incongruent movies; therefore, the hypothesis was supported for the fearful movie condition.

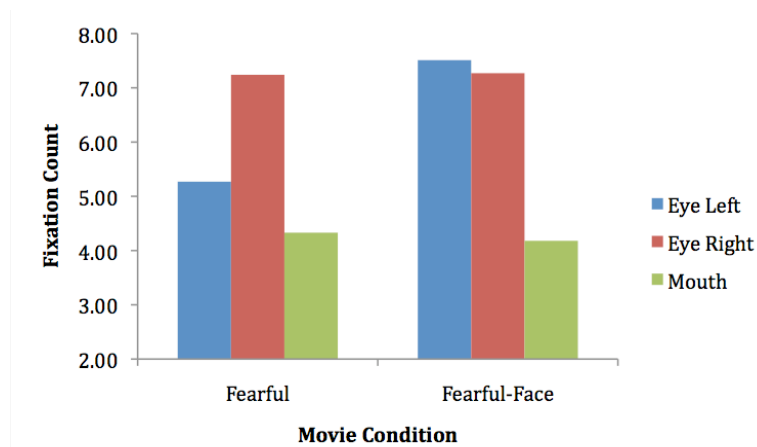


Figure 12. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 2.

.04, $\eta_p^2 = .08$, indicating a significantly greater total fixation duration during incongruent movies when compared to congruent movies, as well as a main effect of Region, $F(2,100) = 8.08$, $p = .001$, $\eta_p^2 = .14$ (Table 12). Contrasts indicated that fixation duration was significantly greater for both the left eye, $F(1,50) = 8.00$, $p = .01$, $\eta_p^2 = .14$, and right eye, $F(1,50) = 15.16$, $p < .001$, $\eta_p^2 = .23$, when compared to the mouth. There was also a significant Movie x Region interaction, $F(1.70,84.86) = 13.49$, $p < .001$, $\eta_p^2 = .21$, indicating that fixation duration for the predefined regions varied by movie type (Figure 12).

Table 13

Descriptive statistics of fixation duration (sec) and count during fearful movies of Task 2

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Fearful	2.58 (2.86)	3.95 (3.09)	1.73 (1.85)	5.27 (5.22)	7.24 (5.09)	4.33 (3.55)
Fearful-Face	3.91 (3.23)	3.48 (2.75)	1.59 (1.89)	7.51 (5.63)	7.27 (5.20)	4.18 (3.73)

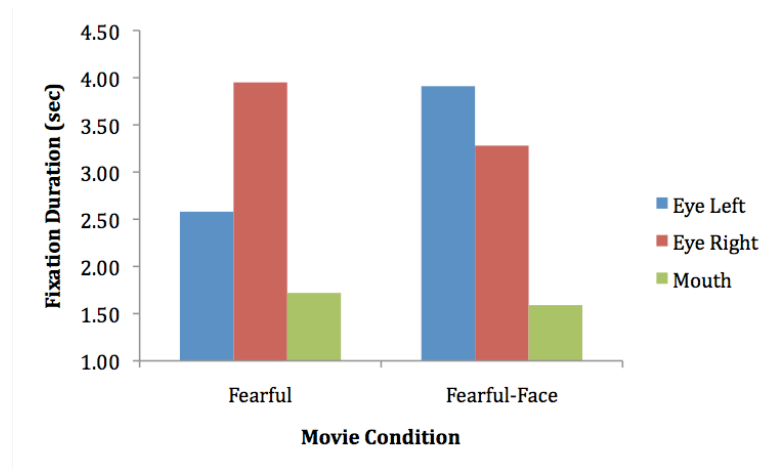


Figure 11. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 2.

Contrasts revealed interactions when comparing congruent and incongruent movies for the left eye compared to the mouth, $F(1,50) = 24.71$, $p < .001$, $\eta_p^2 = .33$. The interaction graph shows that the difference between fixation duration to the left eye and mouth is significantly

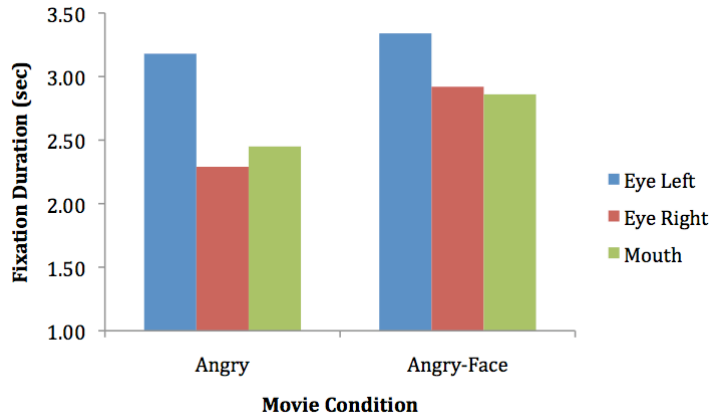


Figure 9. Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 2.

There was a significant main effect of Movie, $F(1,50) = 14.04$, $p < .001$, $\eta_p^2 = .22$, when comparing the *fixation count* for angry and angry-face movies, with a greater number of fixations during incongruent than congruent movies (Figure 11). The main effect of Region and interaction effect were non-significant ($p = .54$ and $p = .06$, respectively). The hypothesis was partially supported by these results; the number of fixations increased significantly for both the eye and mouth regions during the viewing of incongruent movies.

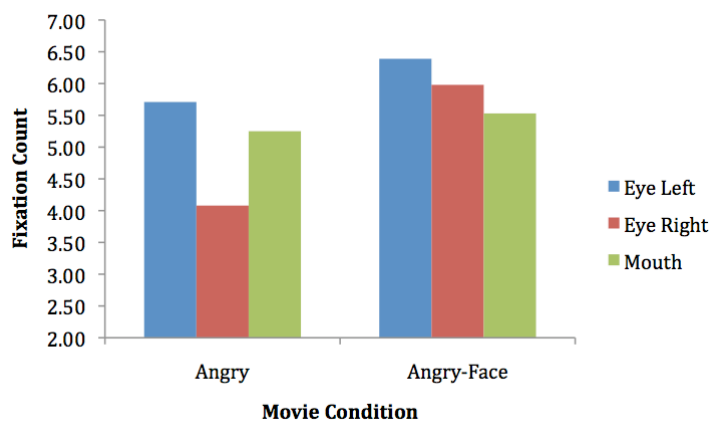


Figure 10. Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 2.

Fearful Movies. Analysis of *total fixation duration* during congruent fearful and incongruent fearful-face movies revealed a significant main effect of Movie, $F(1,50) = 4.28$, $p =$

with this, the previous analyses comparing fixation pattern between congruent and incongruent movies were further explored with the variable of emotion taken into account. Because the addition of emotion as a third variable in the previously described ANOVAs was expected to result in interactions that obscure the primary comparison of interest (movie type), four separate 2 x 3 repeated measures ANOVAs were conducted (i.e., angry vs. angry-face, fearful vs. fearful-face, happy vs. happy-face, neutral vs. neutral-face). As done previously, the left and right eyes were also examined separately in the following exploratory analyses. In order to explicate the results of the following ANOVAs, simple post-hoc contrasts were performed such that both the left eye and the right eye were compared to the mouth.

Angry Movies. The first analysis compared the *total fixation duration* between congruent angry and incongruent angry-face movies (Table 11). Results indicated a significant main effect of Movie, $F(1,50) = 14.29, p < .001, \eta_p^2 = .22$, with a significantly greater fixation duration during the incongruent movies when compared to the congruent movies. The main effect of Region failed to reach significance ($p = .52$). The Movie x Region interaction was not significant ($p = .49$), indicating that there was no differential change in fixation duration to eye and mouth region (Figure 10) between congruent and incongruent angry movies. Therefore, the hypothesis was only partially supported for the angry movie condition; fixation duration increased for both the eye and mouth regions during the incongruent movies.

Table 12
Descriptive statistics of fixation duration (sec) and count during angry movies of Task 2

Emotion	Region					
	Left Eye	Right Eye	Mouth	Left Eye	Right Eye	Mouth
	Duration (sec)			Count		
Angry	3.18 (3.27)	2.29 (2.79)	2.45 (2.40)	5.71 (4.42)	4.08 (4.80)	5.25 (4.69)
Angry-Face	3.34 (3.62)	2.92 (2.74)	2.86 (2.97)	6.39 (5.15)	5.98 (5.14)	5.53 (4.46)

Table 11

Descriptive statistics of fixation duration (sec) and fixation count during congruent and incongruent movies of Task 2

Condition	Region			
	Eyes	Mouth	Eyes	Mouth
	Duration (sec)		Count	
Congruent	24.28 (16.19)	8.57 (8.34)	46.24 (30.41)	18.96 (14.33)
Incongruent	26.52 (17.20)	9.54 (9.62)	52.41 (33.75)	20.73 (15.29)

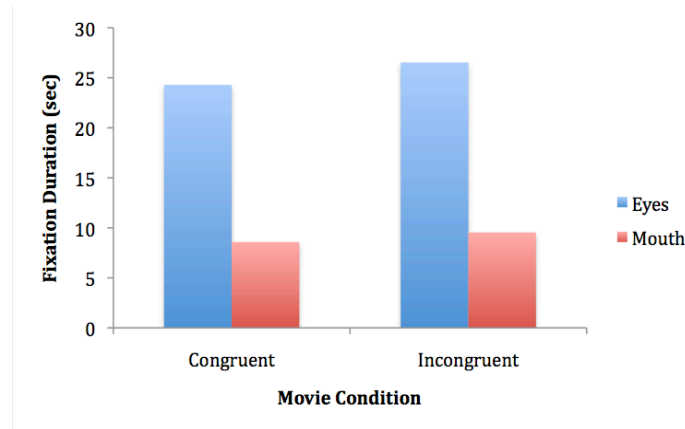


Figure 7. Total fixation duration to the three areas of interest during congruent and incongruent movies of Task 2.

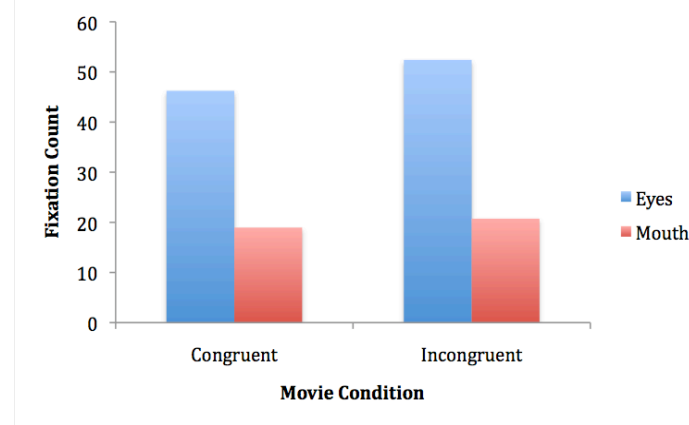


Figure 8. Total fixation count to the three areas of interest during congruent and incongruent movies of Task 2.

Exploratory analyses conducted for Hypothesis 3 (fixation patterns in congruent movies) indicated significant interactions between emotion type and facial region of interest. As a result, it was important to consider the how the variable of emotion may impact the effect of movie condition on *fixation duration* and *fixation count* to the core regions of interest. In accordance

the current hypothesis: eyes and mouth. As done for the Hypothesis 3, each of the two ANOVAs was conducted independently for the dependent variables (DV) *total fixation duration* and *total fixation count*.

When comparing *total fixation duration* between congruent and incongruent movies, analyses revealed a significant main effect of Movie, $F(1,50)=22.54, p<.001, \eta_p^2=.31$ (Table 10). Participants spent a significantly greater amount of time fixating on the regions of interest during the incongruent movies than during congruent movies. There was also a main effect of Region, $F(1,50)=26.05, p<.001, \eta_p^2=.34$, indicating that participants fixated longer on the eye region than the mouth region. The Movie x Region interaction effect ($p=.20$) failed to reach significance (Figure 8). Comparison of *total fixation count* between congruent and incongruent movies again indicated a significant main effect of Movie, $F(1,50)=27.34, p<.001, \eta_p^2=.35$. Participants devoted a significantly greater fixation count to the regions of interest during the incongruent movies than during congruent movies (Figure 9). There was also a main effect of Region, $F(1,50)=27.14, p<.001, \eta_p^2=.35$. Participants devoted a greater fixation count to the eye region than the mouth region. Finally, there was a significant Movie x Region interaction effect, $F(1,50)=7.55, p=.01, \eta_p^2=.13$. This interaction shows that the difference in fixations between the eye and mouth regions is greater during the incongruent movies than the congruent movies. The hypothesis was partially supported by the overall increase in *fixation duration* and *fixation count* during the incongruent movies; however, this increase was not limited to the eye regions. In addition, the increased difference in *fixation count* during the incongruent movies provides partial support for the hypothesis.

during both the congruent and incongruent movies. Similarly, a significantly greater proportion of participants evidenced a featural processing approach during congruent angry, fearful, happy, and neutral movies and the incongruent angry, fearful, happy, and neutral face movies.

Table 10

Proportion of participants utilizing configural or featural processing strategies of Task 2

Condition	Processing Strategy		<i>p</i>
	Configural	Featural	
Task 2 Congruent All Emotions	6	45	<.001
Task 2 Angry	1	50	<.001
Task 2 Fearful	8	43	<.001
Task 2 Happy	4	47	<.001
Task 2 Neutral	4	47	<.001
Task 2 Incongruent All Emotions	12	39	<.001
Task 2 Angry Face	11	40	<.001
Task 2 Fearful Face	9	42	<.001
Task 2 Happy Face	11	40	<.001
Task 2 Neutral Face	9	42	<.001

Hypothesis 6 – Task 2 Fixation Pattern

It was hypothesized that the difference between *total fixation duration* and *total fixation count* devoted to the eye region and the mouth region (i.e., mouth subtracted from eye) would be greater during the viewing of incongruent movies than congruent movies. Specifically, it was predicted that *total fixation duration* and *total fixation count* devoted to the eye region would increase during the viewing of incongruent movies relative to the congruent movies in Task 2. In contrast, the *total fixation duration* and *total fixation count* devoted to the mouth region was expected to decrease during the viewing of incongruent movies relative to congruent movies. The fixation patterns for these two conditions were first compared directly using the congruent movies of Task 2 to the incongruent movies of Task 2 in a series of 2 x 2 (Movie x Region) repeated-measures ANOVAs. The first IV, Movie, consisted of two levels: congruent and incongruent. The second IV, Region, consisted of two levels, the primary regions of interest for

Hypothesis 4 – Task 2 Behavioral Responses and Reaction Time

In order to assess whether individuals demonstrate a bias toward facial emotion when asked to label the emotions presented during the incongruent movie condition, a paired-samples t-test was conducted. The mean number of judgments matching the face ($M=25.49$, $SD=5.04$) was compared to the mean number of judgments matching the voice ($M=4.94$, $SD=4.02$) and revealed a significant bias towards the facially expressed emotion, $t(50)= 16.30$, $p<.001$. A paired samples t-test was utilized to compare response times during the presentation of congruent and incongruent movies in Task 2. Participant response times were significantly longer during incongruent movies ($M=858.97$ ms, $SD=311.06$ ms) than congruent movies ($M=583.10$ ms, $SD=185.41$ ms), $t(50)= -8.99$, $p<.001$.

Hypothesis 5 – Task 2 Processing Approach

To examine the nature of the participants' scanning patterns, differences in facial processing strategies were assessed in the same manner as for the previous hypothesis: the number of participants who used a configural strategy for more than 50% of stimuli in the target condition was compared to the number of participants who used a featural strategy for the majority of trials in a given condition, using separate sign tests for the congruent movies and incongruent movies, as well as for each of the four congruent emotion conditions and the four incongruent movie conditions. For the current hypothesis, that the presentation of incongruent movies would result in an increased reliance on featural processing strategies, the facial processing strategies were compared separately for the congruent and incongruent movies of Task 2. The number of participants exhibiting configural and featural facial processing strategies for each of the experimental conditions can be found in Table 9. Separate signs tests indicated that a significantly greater proportion of participants utilized a featural facial processing strategy

interactions were significant except between the right eye and the nose, again indicating that the number of fixation devoted to these regions was approximately equal across emotions (Figure 7).

Table 9

Within-subjects contrasts for the Emotion \times Region interaction for fixation count to the core regions in Task 1

Contrast			<i>F</i>	<i>p</i>	η_p^2
Emotion	Region				
Angry vs. Fearful	Left Eye	Right Eye	36.12	<.001	.42
		Nose	37.85	<.001	.43
		Mouth	2.78	.10	.05
	Right Eye	Nose	.00	.96	.00
		Mouth	24.95	<.001	.33
		Mouth	Nose	16.54	<.001
Fearful vs. Happy	Left Eye	Right Eye	83.00	<.001	.62
		Nose	60.12	<.001	.55
		Mouth	10.30	.002	.17
	Right Eye	Nose	.87	.36	.02
		Mouth	30.84	<.001	.38
		Mouth	Nose	18.27	<.001
Happy vs Neutral	Left Eye	Right Eye	59.51	<.001	.54
		Nose	43.34	<.001	.46
		Mouth	7.56	.01	.13
	Right Eye	Nose	.32	.58	.01
		Mouth	25.69	<.001	.34
		Mouth	Nose	38.77	<.001

df = 1, 50 for all contrasts

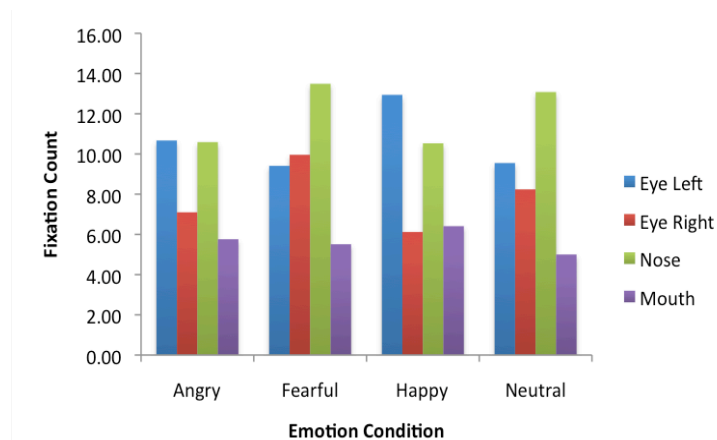


Figure 6. Total fixation count to the four core regions of the face during congruent movies of Task 1.

With regard to differences in *total fixation count* to each of the four core regions of the face during congruent emotion movies, analysis revealed significant main effects of Region, $F(2.05,102.33)= 11.41, p< .001, \eta_p^2= .19$, and of Emotion, $F(3,150)= 7.58, p< .001, \eta_p^2= .13$, (Table 8). Contrasts indicated that fixation count was greater for the left eye than right eye, $F(1,50)= 8.50, p= .01, \eta_p^2= .15$, and the mouth, $F(1,50)= 13.70, p=.001, \eta_p^2= .22$, greater for the nose than the right eye, $F(1,50)= 9.58, p=.003, \eta_p^2= .16$, and greater for the right eye than the mouth, $F(1,50)= 4.14, p= .05, \eta_p^2= .08$. In addition, contrasts showed that fixation count was significantly greater during fearful than both angry movies, $F(1,50)= 25.76, p< .001, \eta_p^2= .34$, and happy movies, $F(1,50)= 7.02, p= .01, \eta_p^2= .12$.

Table 8

Descriptive statistics of fixation count during congruent movies of Task 1

Emotion	Region			
	Left Eye	Right Eye	Nose	Mouth
Angry	10.76 (7.01)	7.10 (6.08)	10.59 (6.03)	5.76 (4.55)
Fearful	9.41 (6.21)	9.96 (6.51)	13.49 (7.03)	5.51 (4.49)
Happy	12.94 (7.21)	6.12 (5.74)	10.53 (6.03)	6.41 (5.46)
Neutral	9.55 (7.47)	8.24 (6.01)	13.08 (6.23)	5.00 (4.31)
Total	42.67 (26.35)	31.41 (22.64)	47.69 (23.65)	22.69 (17.16)

There was also a significant Region x Emotion interaction for *total fixation count*, $F(7.35,367.35)= 17.91, p< .001, \eta_p^2= .13$, which indicated that the number of fixations for each of the core regions varied significantly by emotion presented. Contrasts (Table 9) revealed a number of significant interactions. When comparing the angry and fearful movies, significant interactions exist for all regions except between the left eye and mouth and the right eye and the nose, indicating that the difference between these regions is comparable for both emotion conditions. When comparing fearful and happy movies and happy and neutral movies, all

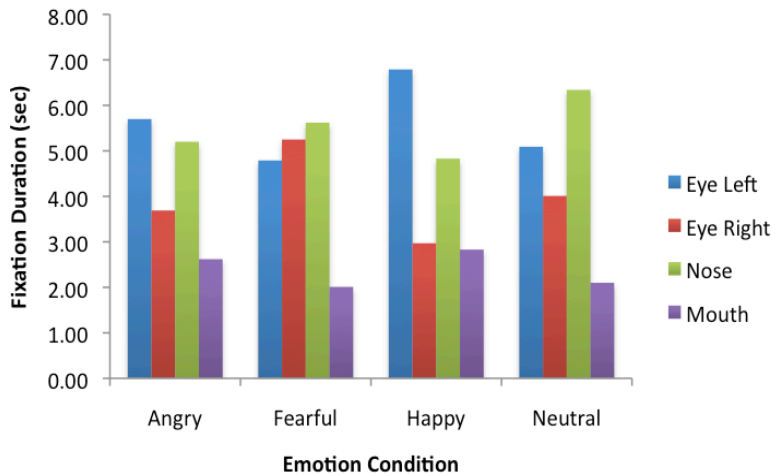


Figure 5. Total fixation duration to the four core regions of the face during congruent movies of Task 1.

Table 7

Within-subjects contrasts for the Emotion \times Region interaction for fixation duration (sec) to the core regions in Task 1

Contrast			<i>F</i>	<i>p</i>	η_p^2
Emotion	Region				
Angry vs. Fearful	Left Eye	Right Eye	17.80	<.001	.26
		Nose	10.10	.003	.17
		Mouth	.54	.47	.01
	Right Eye	Nose	4.81	.03	.09
		Mouth	25.70	<.001	.34
		Mouth	Nose	6.41	.02
Fearful vs. Happy	Left Eye	Right Eye	52.71	<.001	.51
		Nose	33.05	<.001	.40
		Mouth	6.80	.01	.12
	Right Eye	Nose	6.48	.01	.12
		Mouth	39.31	<.001	.44
		Mouth	Nose	10.80	.002
Happy vs Neutral	Left Eye	Right Eye	43.61	<.001	.47
		Nose	32.20	<.001	.39
		Mouth	6.62	.01	.12
	Right Eye	Nose	1.31	.26	.03
		Mouth	25.42	<.001	.34
		Mouth	Nose	29.31	<.001

df = 1, 50 for all contrasts

approached significance for the right eye compared to the nose ($p=.06$). There was no main effect of Emotion ($p=.32$).

Table 6

Descriptive statistics of fixation duration (sec) for congruent emotion movies of Task 1

Emotion	Region			
	Left Eye	Right Eye	Nose	Mouth
Angry	5.70 (4.14)	3.69 (3.23)	5.20 (3.45)	2.62 (2.77)
Fearful	4.79 (3.70)	5.25 (3.82)	5.62 (3.74)	2.00 (2.18)
Happy	6.79 (4.55)	2.97 (2.95)	4.83 (3.41)	2.83 (2.78)
Neutral	5.09 (4.30)	4.01 (3.19)	6.34 (3.98)	2.10 (2.26)
Total	22.37 (15.90)	15.92 (12.17)	21.99 (13.61)	9.56 (9.26)

There was also a significant Region x Emotion interaction for total fixation duration during congruent emotion movies, $F(6.61, 330.64) = 14.78$, $p < .001$, $\eta_p^2 = .23$. This Region x Emotion interaction is presented in Figure 6 and indicates that the duration of fixations for each of the core regions described previously varied significantly by emotion presented. Contrasts (Table 7) revealed a number of significant interactions. When comparing the angry and fearful movies, significant interactions exist between all regions except the left eye and mouth, indicating that the difference between these two regions is comparable for both emotion conditions. Interactions were found between all core regions when comparing fearful and happy movies. Finally, comparing happy and neutral movies revealed significant interactions between all regions except the right eye and the nose; the difference between these two regions is similar for both movie conditions.

$F(1,50)=14.94, p<.001, \eta_p^2 = .23$, and the mouth, $F(1,50)=56.19, p<.001, \eta_p^2 = .53$. Similarly, the one-way repeated-measures ANOVA examining *total fixation count* to the three core regions revealed a significant effect of Region, $F(1.28,63.91)=33.35, p<.001, \eta_p^2 = .40$. Contrasts revealed that participants fixated more frequently on the eye region than both the nose, $F(1,50)=11.98, p=.001, \eta_p^2 = .19$, and the mouth, $F(1,50)=52.27, p<.001, \eta_p^2 = .51$. These findings are in support of the hypothesis that the eye region is fixated on for longer and more often than other core regions of the face.

Table 5

Descriptive statistics of fixation duration (sec) and fixation count during congruent movies of Task 1

	Region					
	Eyes	Nose	Mouth	Eyes	Nose	Mouth
Condition	Duration (sec)			Count		
Congruent	38.29 (20.26)	21.99 (13.61)	9.56 (9.26)	74.08 (40.67)	47.69 (23.65)	22.69 (17.16)

Analysis revealed that participants showed a bias toward one eye (left or right) in a subset of experimental conditions. In order to examine these difference in the context of the other core regions and to explore the pattern of fixations across the four emotions conditions, two 4x4 (Region x Emotion) repeated-measures ANOVAs were performed. The first IV in this ANOVA, Region, now consisted of four levels: left eye, right eye, nose, and mouth (Table 6), and the second IV, Emotion, consisted of the four emotion conditions. With regard to *total fixation duration*, analysis revealed a significant main effect of Region, $F(2.38,118.78)= 8.88, p<.001, \eta_p^2 = .15$. Contrasts revealed that fixation duration was significantly greater for the left eye than the right eye, $F(1,50)= 5.42, p= .02, \eta_p^2 = .10$, and the mouth, $F(1,50)= 16.42, p<.001, \eta_p^2 = .25$. Fixation duration failed to reach significance when comparing the left eye and nose ($p=.91$), but

with significantly greater fixation count in the core regions as compared to the periphery.

Analysis revealed a significant main effect of Emotion, $F(3,150)=3.86$, $p=.01$, $\eta_p^2=.07$.

Contrasts indicated that fixation count was greater for fearful than for angry stimuli, $F(1,50)=10.52$, $p=.002$, $\eta_p^2=.17$. There was also a significant Region x Emotion interaction, $F(3,150)=7.83$, $p<.001$, $\eta_p^2=.14$. Contrasts indicated that core fixation count, when compared to peripheral fixation count, differed between angry and fearful, $F(1,50)=25.70$, $p<.001$, $\eta_p^2=.34$, and between fearful and happy, $F(1,50)=12.10$, $p=.001$, $\eta_p^2=.20$. Looking at the interaction graph (Figure 5), these effects reflect the fact that the difference between fixation count to the core versus the periphery was significantly greater for fearful stimuli than for both angry and happy stimuli.

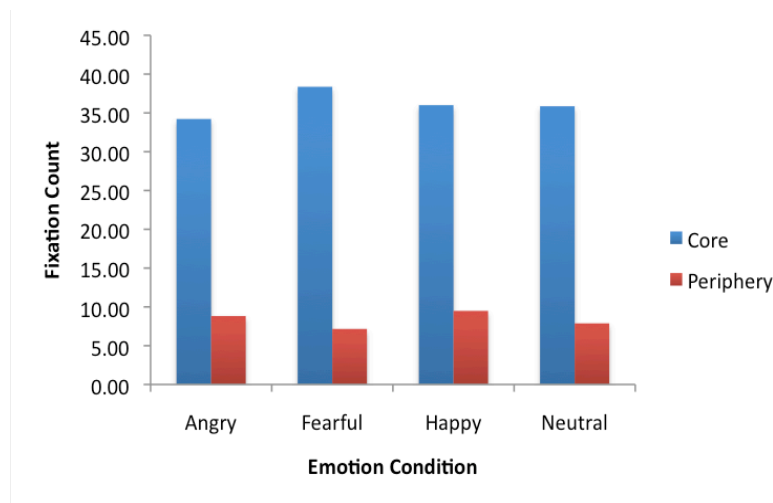


Figure 4. Total fixation count to the core and peripheral regions during congruent movies of Task 1.

Core Regions. In order to examine the differences in *total fixation duration* to each of the three core regions of the face during congruent movies, a one-way repeated measures ANOVA was conducted. The IV, Region, consisted of three levels: eyes, nose, and mouth (Table 5). There was significant effect of Region, $F(1.38,68.99)=33.94$, $p<.001$, $\eta_p^2=.40$. Contrasts indicated that participants fixated for longer periods of time on the eye region than both the nose,

In order to examine the pattern of fixations across the four emotion conditions, two 4x2 (Region x Emotion) repeated-measures ANOVAs were performed. The first independent variable (IV), Region, consisted of two levels: core and periphery. The second IV, Emotion, consisted of four levels: angry, fearful, happy, and neutral. Mean fixation durations can be seen in Table 4. The repeated-measures ANOVA examining *total fixation duration* revealed a significant main effect of Region, $F(1,50) = 508.30, p < .001, \eta_p^2 = .91$, with fixation duration significantly greater for the core region than the periphery. The main effect of Emotion was not significant ($p = .10$). The significant Region x Emotion interaction, $F(3,150) = 2.89, p = .04, \eta_p^2 = .06$, indicates that fixation duration to each of the core regions of the face varies by the type of emotion portrayed in the stimulus. Contrasts revealed differences in fixation duration when comparing the core to the periphery for angry compared to fearful, $F(1,50) = 7.23, p = .01, \eta_p^2 = .13$. This interaction indicates that fixation preference for the core regions of the face, when compared to the periphery, was significantly more pronounced for fearful stimuli than angry stimuli (Figure 4).

Table 4
Descriptive statistics for fixation data during congruent movies of Task 1

Emotion	Region			
	Core	Periphery	Core	Periphery
	Duration (sec)		Count	
Angry	17.20 (3.57)	2.74 (2.16)	34.22 (10.33)	8.82 (7.22)
Fearful	17.67 (3.39)	1.99 (1.88)	38.37 (10.12)	7.16 (6.72)
Happy	17.42 (3.15)	2.62 (1.88)	36.00 (11.06)	9.49 (6.13)
Neutral	17.54 (3.75)	2.24 (2.14)	35.86 (9.66)	7.88 (7.18)

A second repeated-measures ANOVA was conducted to examine differences in *total fixation count* for the core regions and the periphery during the four congruent emotion conditions. There was a significant main effect of Region, $F(1,50) = 275.37, p < .001, \eta_p^2 = .85$,

and then for the left and right eye separately. Each of the t-tests and subsequently described ANOVAs was conducted independently for two dependent variables (DV): *total fixation duration* and *total fixation count*.

In order to examine differences revealed by the second set of ANOVAs comparing the core regions of the face, both simple and repeated contrasts were performed for the independent variable of Region to allow comparison of predefined regions critical to the hypotheses. Simple contrasts were performed for the ANOVA comparing eyes, nose and mouth, allowing for the comparison of the eye region to the other two core regions. When comparing the left eye, right eye, nose, and mouth, simple contrasts allowed for comparison of the left eye region to each of the other core regions (right eye, nose, and mouth) and for comparison of the right eye to the mouth. Repeated contrasts allowed for the comparison of the right eye to the nose. Repeated contrasts were performed for the independent variable of Emotion.

Core vs. Periphery. To determine whether participants devoted a greater *fixation duration* and *fixation count* to the core regions of the face than to the periphery during congruent movies, paired-samples t-tests were conducted. With regard to *total fixation duration*, results indicated that participants fixated significantly longer on the core regions of the face ($M=69.84$ sec, $SD=13.25$) than the peripheral regions ($M=9.59$ sec, $SD=7.16$) during the congruent movies, $t(50)=22.55, p<.001$. Similarly, participants devoted a significantly greater *fixation count* to the core regions ($M=144.45$, $SD=38.25$) than to the peripheral regions ($M=33.35$, $SD=24.64$) during the congruent movies, $t(50)=16.59, p<.001$. These findings are in support of the hypothesis that participants would fixate significantly for longer and more often on the core than peripheral regions.

processing approaches, comparing participants who most often used a configural strategy to those who most often used a featural strategy, fixations to the left and right eyes were combined to yield an overall eye region measure. The number of participants who used configural and featural processing strategies for each of the experimental conditions can be found in Table 3. Separate signs tests indicated that a significantly greater number of participants utilized a featural processing strategy overall and in every emotion condition. These findings failed to support the hypothesis that participants would utilize a configural processing strategy during an emotion-decision task when viewing congruent movies.

Table 3
Proportion of participants utilizing configural or featural processing strategies of Task 1

Condition	Processing Strategy		<i>p</i>
	Configural	Featural	
Task 1 All Emotions	10	41	<.001
Task 1 Angry	6	45	<.001
Task 1 Fearful	9	42	<.001
Task 1 Happy	17	34	.02
Task 1 Neutral	7	44	<.001

Hypothesis 3 – Task 1 Fixation Pattern

The first analyses examined differences in fixation patterns between the core regions and peripheral regions of the face. Paired-samples t-tests were conducted to explore fixation patterns during all congruent movies and 4x2 repeated-measures ANOVAs explored these differences across the four congruent emotion conditions. Following these analyses, a series of repeated-measures ANOVAs were conducted to evaluate differences in fixation patterns for the core regions of interest of the face during all congruent movies and across the four different congruent emotion conditions. As a result of the significant differences between fixation duration to the left and right eye presented previously, analyses were initially compared for the eye region combined

Results indicated that participants spent a significantly greater amount of time fixating on the left eye than the right eye during the congruent happy movies of Tasks 1 2, and 3, the incongruent happy-face movies of Task 2, and the congruent angry movies of Task 1. In contrast, participants spent a significantly greater amount of time fixating on the right eye during the viewing of congruent fearful movies in Tasks 2 and 3. As a result of these biases, all fixation pattern analyses (Hypotheses 3, 6, and 9) were conducted for left and right eyes separately. Analyses conducted for processing approach (Hypotheses 2, 5, and 8) were conducted with the left and right eyes combined, as the entire eye region was of primary interest.

Hypothesis 1 – Task 1 Behavioral Responses

In order to determine whether the participants, when viewing congruent movies, were able to identify correctly the emotion portrayed by the actors, the percentage of correct responses was calculated across all congruent movies of Task 1 (n=40). Results indicated that participants responded correctly to 97.50% of the congruent movies.

Hypothesis 2 – Task 1 Processing Approach

To examine the nature of the participants' scanning patterns, differences in regions and fixation patterns were assessed for the entire task, and separately for each of the four emotion conditions. The number of participants who predominantly used configural strategies, fixating on both the eye region and the mouth region during more than 50% of the movies in one condition, was compared to the number of participants who used a featural strategy, limiting their fixations to either the eye or the mouth region during the majority of movies in a particular condition, using a sign test. As previously stated, the eye region in its entirety and the mouth region, indicating features to both the upper and lower halves of the face, were the regions of interest for analyses of processing approach. As such, for this and all subsequent analyses of facial

eye). This naming convention follows from the fact that several researchers have explored the lateralization of emotion perception and refer to the advantage of presentation to the left or right visual field or the preferred sampling of the left or right side of a stimulus (Asthana & Mandal, 2001; Borod, 1988; Marcus, 2005). The hemi-space of the stimulus is consistently labeled (left or right) from the perspective of the viewer in order to make inferences about hemispheric activation or involvement.

Prior to conducting analyses for the three hypotheses related to fixation pattern across each of the predefined core regions, the total fixation duration devoted to the left and right eye was compared using paired-samples t-tests for each of twenty experimental conditions (Table 2).

Table 2

Contrast between fixation duration for the left eye and right eye

Condition	Fixation Duration (sec)				<i>t</i> (50)	<i>p</i>
	Left Eye		Right Eye			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Task 1 Angry	5.70	4.15	3.69	3.23	2.68	.01
Task 1 Fearful	4.79	3.70	5.25	3.82	-.61	.55
Task 1 Happy	6.79	4.55	2.97	2.95	4.72	<.001
Task 1 Neutral	5.09	4.30	4.01	3.19	1.50	.14
Task 2 Angry	3.18	3.27	2.29	2.79	1.43	.16
Task 2 Fearful	2.58	2.86	3.95	3.09	-2.44	.02
Task 2 Happy	3.53	3.72	2.15	2.31	2.18	.03
Task 2 Neutral	3.92	3.76	2.89	2.45	1.58	.12
Task 2 Angry Face	3.34	3.62	2.92	2.74	.64	.53
Task 2 Fearful Face	3.91	3.23	3.48	2.75	.77	.45
Task 2 Happy Face	4.13	3.98	2.53	2.11	2.47	.02
Task 2 Neutral Face	3.66	3.87	2.58	2.70	1.65	.10
Task 3 Angry	3.47	3.78	2.43	2.64	1.69	.10
Task 3 Fearful	2.75	2.79	4.32	3.31	-2.74	.01
Task 3 Happy	3.59	3.48	2.08	2.29	2.59	.01
Task 3 Neutral	2.82	3.39	2.09	2.26	1.44	.16
Task 3 Angry Face	2.80	2.96	2.42	2.72	-.00	.99
Task 3 Fearful Face	2.42	2.73	3.42	3.20	-1.78	.08
Task 3 Happy Face	3.22	3.64	2.67	2.73	.84	.40
Task 3 Neutral Face	2.82	3.04	2.32	2.68	.88	.38

always indicated “Yes” and the right button always indicated “No.” The participant was instructed to respond within a 5-second response period.

These three tasks were administered in the same order for all participants; randomized presentation of the tasks would have contaminated *Tasks 1* and *2*. Because the initial task examined an individual’s scanning and fixation pattern when congruent movies were presented without exposure to incongruent movies, *Tasks 2* and *3* must follow *Task 1*. In addition, premature exposure to the question “Does the emotion in the face match the emotion in voice?” would have cued participants to attend to each modality separately. Cueing participants to the fact that facial emotion and prosody did not always match one another would have influenced their emotion judgments and thus the third task could not be presented prior to the first two tasks.

Results

Mauchly’s test was considered for all within-subjects analyses. For all effects in which the assumption of sphericity had been violated, degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (See Appendix 2 for all data regarding Mauchly’s test).

For each of the following analyses which compare processing approach or fixation pattern between emotion conditions, incongruent movies were grouped for analysis by the emotion expressed in the actor’s face. For example, a happy-face movie refers to one in which the facial expression is that of happy and the vocal emotion is not happy, meaning that it is either angry, fearful, or neutral. In contrast, a happy movie refers to a congruent movie in which both face and voice are happy.

In addition, it should be noted that the left eye refers to the left side of the stimulus (i.e., the actor’s right eye) and the right eye refers to the right side of the stimulus (i.e., the actor’s left

male actor delivered a happy-congruent sentence in one movie (i.e., happy face/happy voice) and four of the incongruent movies included the male actor delivering any four sentences with a happy facial expression (i.e., happy face/angry voice, happy face/fearful voice). After each movie, the participant was asked to respond, in a forced-choice format, with one of the four emotion choices: “Happy”, “Angry”, “Fearful”, and “Neutral”. The response procedure was the same as in *Task 1*, with only four labeled keys available and the instructions to respond within a 5-second response period. Because of the forced-choice format of this study, individuals were not afforded the opportunity to respond with a blended emotion. The response choice indicated which of the two modalities had a stronger influence on emotion perception.

Task 3. In the third task, the participants were shown both congruent and incongruent emotion movies and asked to decide if the facial expression and prosody match one another by answering the question, “Does the emotion in the face match the emotion in the voice?” The participants were shown a total of 64 movies (32 congruent/32 incongruent). Movie selection proceeded as described for *Task 2*, but movies selected for *Task 3* were primarily chosen from those not included in *Task 2*. Because *Task 2* and *Task 3* consisted of 32 and *Task 1* of 40 congruent movies, some congruent sentence and emotion combinations were repeated (n=27). Selection of congruent movies for *Task 2* and *Task 3* occurred independently of those previously chosen; no overlap between *Task 2* and *Task 3* occurred and overlap between these and *Task 1* was kept to a minimum while ensuring that emotion and sentence combinations were distributed appropriately. As with previous tasks, this task employed a forced choice format because the participant was again unable to respond verbally. The participant was instructed to respond by pressing the right or left button on the mouse to indicate his or her response: the left button

Task 1. In the first task, the participants were shown 40 congruent movies and asked, “How does this person feel?” Movies were selected to include ten movies portraying each of four emotions. Half of each emotion set consisted of sentences delivered by the male actor and the other half by the female actor, with randomized selection of the sentences delivered by each. Because the participants were unable to respond verbally, these tasks employed a forced choice response format. After each movie, a screen was displayed that provided four emotion choices: “Happy”, “Angry”, “Fearful”, and “Neutral.” The individual was seated in front of a keyboard on which only four labeled keys were available (i.e., S, C, N, K). Each emotion word cue corresponded spatially to the key that the participant was to use to respond with that emotion (i.e., the words and keys corresponded to each other from left to right). The participant was instructed to respond within the 5-second response period provided after each movie; the task automatically proceeded to the next movie after a response was made or at the end of this 5-second response period. The remaining keys were covered with a rubber cover. The participant’s hands were positioned so that he/she could rest four fingers (e.g., index and middle fingers) comfortably on the operable keys.

Task 2. In the second task, participants were shown both congruent and incongruent movies and asked, “How does this person feel?” The participant was shown a total of 64 movies (32 congruent/32 incongruent). Half of each movie type consisted of sentences delivered by the male actor, the other half by the female actor. For the congruent movies, each actor delivered four movies in each of the four emotions, with sentences chosen at random. For the incongruent movies, sentences were again chosen at random. Incongruent movies were selected to ensure that both actors facially express each of the four emotions four times. In the end, each actor facially expressed each emotion eight times (i.e., four congruent, four incongruent). For example, the

The second part of the session consisted of the non-invasive eye-tracking of participants while they viewed the emotional movie stimuli. Participants were allowed to acclimate to sitting in the chair for as long as required for their comfort. The PI guided each participant through a practice run, watching non-experimental stimuli and responding with the four designated keys on the keyboard. The eye-tracking apparatus was then be calibrated before beginning the experiment. Each participant was administered three experimental movie tasks. As previously described, gaze data was also collected during the administration of the DANVA2 Adult Facial Expressions subtest for 10 participants. Instructions for each task were provided before each task began. The total viewing time was approximately 30-45 minutes, for a total session time of 2-2.5 hours.

In this study, the participants were shown DAVE stimuli one at a time and, during a five second response period, asked to make a response. If the participant responded prior to the five-second limit, the task advanced automatically. Stimuli were presented using DirectRT (Empirisoft Corporation, 2006) on a Dell Intel ® Pentium ® 4 CPU desktop computer using Windows XP operating system. Each stimulus was presented at 500x420 pixels. DirectRT recorded the participant's behavioral responses (emotion identification, match/mismatch judgment) and response time. Participants were seated approximately 60-cm from the computer screen and stimuli subtended a visual angle of approximately 36° vertical x 30° horizontal. Before beginning the eye-tracking tasks, participants were provided with general instructions. Prior to each of the three individual tasks, participants were instructed to "Pay attention to the movie" and provided with the specific instructions for that task. The script used for delivery of instructions can be found in Appendix 1.

fixation duration across all movies within each of the 20 experimental conditions: conditions were task (1,2,3), movie type (congruent, incongruent), and emotion (angry, fearful, happy, neutral) specific. For example, the number of seconds spent fixating on the left eye during each of the ten angry movies in Task 1 was summed, yielding the total fixation duration for the left eye during angry congruent movies of Task 1. Similarly, total fixation count for each region was computed by summing the number of fixations lasting 200 ms or longer for each of the movies in one experimental condition, such as angry congruent.



Figure 3. Sample of male and female stimuli with defined facial regions.

Experimental Procedure. The current study consisted of two parts, neuropsychological testing and the experimental tasks, and administration order was counterbalanced across participants. The first part of the session consisted of neuropsychological testing and screening for all potential participants. These assessments included the previously described cognitive, perceptual, and emotional tests and lasted approximately 1-1.5 hours. Participants were excluded if they had a perceptual disorder that would interfere with the detection of stimuli or if they indicated a risk of meeting DSM-IV-TR criteria at the time of screening for a form of major psychopathology (e.g., depression, anxiety, ASD). Individuals were offered periodic breaks and were told that they may request a break at any time during testing.

assigned to participants in the order of enrollment. The signed consent forms served as the only paperwork that connected an individual participant's name to his or her assigned tag; the consent forms were filed in a cabinet separate from the data collected and accessible only to researchers. Computers housing any information pertaining to the study were password-protected and locked in the research lab. Any files that were printed were stored in a locked filing cabinet, accessible only to researchers.

Data Analysis Procedure. A fixation was defined as a set of consecutive data points during which the point of gaze (POG) had a standard deviation of no more than 0.5 degrees, in either the horizontal or vertical plane, from the original POG. Each fixation began when six sequential gaze positions met these criteria, a period of approximately 102 ms, and ended when the number of sequential gaze positions deviated from the original POG by a degree angle of one or more in either the horizontal or vertical plane. For the current study, spatial data were compared in four areas of interest, defined individually for each movie and based on the location of core facial features (i.e. eyes, nose, mouth). A static image, or one frame of each movie clip, was used to define these regions. Faces were divided into 4 blocks that covered these three core facial regions, with separate blocks for right and left eye. An example of the defined facial regions can be seen in Figure 3. The remaining blocks were considered as the region peripheral to the face (Marcus, 2005; Pelphrey, 2002).

Number of fixations and fixation duration were calculated separately for each of the four predefined regions of the face: left eye, right eye, nose, and mouth. The number of fixations and fixation duration for the left eye, right eye, nose, and mouth (core) were also combined to allow for comparison to number of fixations and fixation duration for all other regions of the stimuli (peripheral region). Total fixation duration for each region was computed by summing the total

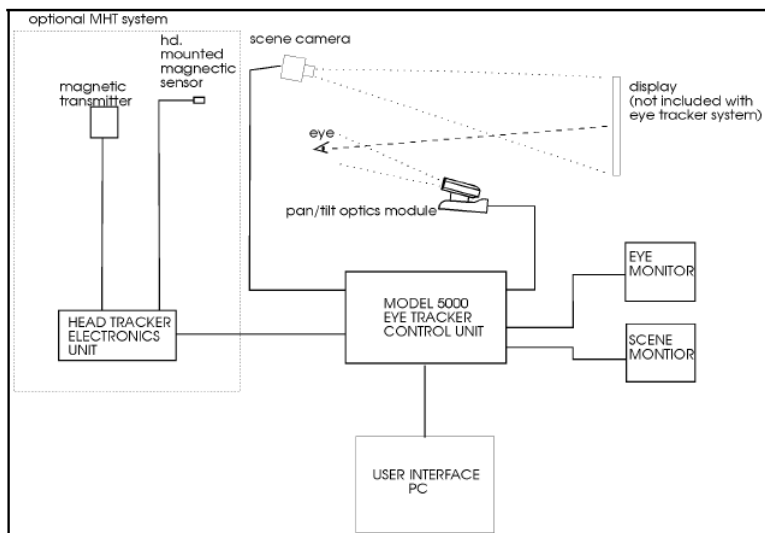


Figure 2. Schematic of ASL 6000 set-up and eye-tracking data collection.

Procedure

Participants were recruited from the undergraduate participant pool at Georgia State University using an advertisement on the Sona Experiment Managements System (SONA). For their participation in the study, students of Georgia State University received three credits for their undergraduate course in Psychology. All tasks took place in a research laboratory in the Psychology Department of Georgia State University, Atlanta, Georgia. All individuals participated in testing during one session, with informed consent procedures at the start of the session. Upon arrival at the Psychology Department, the principal investigator explained the entire study and procedures to the participant. The consent form was reviewed verbally by the PI and was signed by the participant. The PI addressed any concerns before obtaining written consent.

The researchers used several precautions to ensure the confidentiality of information obtained during the course of this study. In no case did the participants' actual name or other identifying information appear on any computerized or written test form or on any other data files created in this experiment. Each participant was identified using a unique alphanumeric tag,

system uses a table-mounted chin-rest to eliminate head movement during recording. The eye-tracker camera produces a near infrared light beam that shines into the eye. A portion of that light is reflected off of the retina through the pupil and back to the camera, which rests in front of the presentation computer. The camera uses this reflected light to localize both the participant's pupil and an image reflection on the cornea. A schematic can be viewed in Figure 2. By computing the vector between the corneal reflection (a fixed image) and the papillary position (an indicator of direction of gaze), the camera can identify and record the participant's point of gaze (POG). Points of gaze were collected and recorded by EYENAL, the data analysis software, at a rate of 60 samples per second, providing 60-127.8 data points for each stimulus (stimulus length ranges from 1.00 to 2.13 seconds). Eye blinks by the participant or loss of the corneal reflection, identified by a loss of data from the pupil-corneal camera, were excluded from final analysis. A blink was defined as fewer than 12 consecutive lost data fields, where the pupil diameter equaled zero. When a blink occurred during a fixation, then the loss of data did not cause the fixation to end. Consecutive lost fields, where the pupil diameter equaled zero, greater than or equal to 12 were considered to be losses. When losses occurred during a fixation, the fixation measurement ended. When a loss of corneal reflection was detected, despite valid pupil data, the recorded values were interpreted as invalid data.

Prior to the start of each movie, a white cross-hair appeared in the center of the screen to ensure that all perceivers began viewing each stimulus from the same point (Manor et al., 1999; Marcus, 2005). Data was collected continuously throughout each of the three tasks; external data signals were sent from DirectRT to EYENAL to mark the beginning and end of each movie stimulus. EYENAL software provided both spatial and temporal eye gaze data for analysis. Spatial data consisted of the location of fixation points on each DAVE stimulus.

were excluded from the current study due to elevated error scores, as indicated by a mean number of errors greater than or equal to two standard deviations below the norm.

For the current study, the Adult Paralanguage subtest was administered in its original format. The Adult Facial Expression subtest was administered using DirectRT, in the same manner previously described for the three experimental tasks, for a subset of participants (n=10). This manner of presentation allowed for the collection of gaze data during the presentation of static emotion stimuli, providing a validated comparison for the DAVE stimuli; such analysis is beyond the scope of the current study. The remaining participants (n=41) were administered the Adult Facial Expression subtest in the original format.

Adult Paralanguage (DANVA2-AP). The Adult Paralanguage subtest of the DANVA2 consists of male and female actors stating the sentence “I’m going out of the room now, but I’ll be back later.” After the presentation of this recorded sentence, the participant is asked to indicate the emotion portrayed in the voice by selecting from one of four choices: happy, sad, angry, and fearful. Sentences are played one at a time and can be repeated as often as the participant requires.

Adult Facial Expressions (DANVA2-AF). The Adult Facial Expressions subtest of the DANVA2 consists of 24 photographs of male and female adults displaying one of four facial expressions: happy, sad, angry, and fearful (Nowicki & Carton, 1993). During administration of this subtest, photographs are displayed one at a time, for 2.5-seconds each. Participants are asked to identify the emotion portrayed in the face.

Apparatus

Eye-tracking data was collected using a bright pupil image technique and hardware and software created by Applied Science Laboratories[®] (ASL; Bedford, MA). The ASL 6000 tracking

verbal auditory discrimination, the reaction time is highly sensitive to the participant's ability to attend to and concentrate on the task. Test-retest differences are minimal and internal reliabilities (split-half and odd-even) are reported as .77 and .62, respectively. This test was administered to ensure that final participants possessed within normal abilities of auditory processing.

Participants whose data indicated a mean number of errors greater than or equal to two standard deviations below the mean of a normative sample (i.e., score ≤ 20) were to be excluded from the current study; no participants were excluded on the basis of this criterion.

Diagnostic Analysis of Nonverbal Accuracy, Second Edition (DANVA2; Nowicki, 2004). In order to assess the participants' ability to detect nonverbal cues, participants underwent testing on two of the four subtests of the DANVA2. Each subtest includes 24 trials and asks the viewer to identify, in a forced-choice format, the facial expression depicted in the photograph (Adult Facial Expressions) or the emotion portrayed in the actor's (both male and female) voice (Adult Paralanguage). Two additional subtests, Child Facial Expressions and Child Paralanguage, were not administered in the current study. On each of the subtests, the participants are given the same four response choices (happy, sad, angry, and fearful). The DANVA2 is a computer-based test and the responses are recorded according to which of the four emotion labels are selected with the mouse. Results were found to be consistent over time, with test-retest reliability ranging from .81 to .84 for facial expressions and .73 to .93 for paralanguage. Internal consistency was good, with coefficient alpha ranging from .78 to .83. Criteria related validity has been shown to be good, such that scores on the DANVA2 correlated closely with scores on various psychological measures. See Nowicki & Duke (1994) and Nowicki (2004) for further information on psychometrics and scale creation. These subtests were administered to assess an individual's ability to decode nonverbal information. No participants

ensure complete understanding of the tasks, individuals who scored lower than 70 on this measure were to be excluded from the current study; no participants were excluded on the basis of this criterion.

Benton Facial Recognition Test (BFRT; Benton, 1994). The BFRT assesses an individual's ability to accurately recognize human faces without requiring recall. Each participant was exposed to three different conditions in which they are asked to correctly match human faces: matching of identical front views, matching of a front view with a three-quarter view, and matching a front views under two different lighting conditions. On the long-form of the BFRT, there are 54 separate matches, using a total of 22 stimulus cards. The short-form, from which a total-score is extrapolated, consists of 27 separate matches and uses a total of 13 stimulus cards. The short-form of the BFRT was administered for the current study. On the short-form, there are six items in which the participant is asked to select the single stimulus card that correctly matches a sample photograph. There are seven items in which the participant is asked to select the three stimulus cards that correctly match a sample photograph. The test is commonly used by neuropsychologists and the authors report extensive normative data for individuals over age 16 from various studies indicating good reliability and validity (Benton et al., 1983). The test was used to screen out individuals who did not possess typical facial processing abilities. None of the participants of the current study were excluded due to impaired facial processing abilities (i.e., score \leq 38).

Seashore Rhythm Test (Seashore, Lewis, & Sævi, 1960). This task requires participants to determine if two rhythms presented are matching or mismatching. The task is taken from a subtest of the Halstead-Reitan Neuropsychological Test Battery and requires participants to discriminate between 30 pairs of rhythmic beats. Classified as a measure of non-

whether participants had significant behavioral or psychological problems. Because of previous research findings that symptoms of depression, anxiety, or psychosis may bias or impede an individual's perception of emotions, those participants exhibiting clinically significant symptoms on the Depression and/or Anxiety scales, as indicated by T-scores of ≥ 70 , were excluded from the current study. A total of six participants were excluded from the current study due to clinically significant symptom of anxiety ($n=2$), depression ($n=3$), or both anxiety and depression ($n=1$). Those participants who self-report significant symptoms of psychopathology were provided with information about and referred to the Georgia State University Counseling Center.

Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Participants were administered the WASI to estimate overall cognitive ability. The WASI consists of four subtests (Vocabulary, Block Design, Similarities, and Matrix Reasoning) used to assess various aspects of intelligence, including Full Scale IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ). These four subtests were selected for the abbreviated scale because they possess the highest loadings (i.e. $> .70$) on g, or general intelligence, and this maximizes the correlation between the WISC-IV (FSIQ) and the WAIS-III FSIQ. Furthermore, the reliability of these four subtests has been shown to be exceptionally high (Wechsler, 1999). The WASI is a psychometrically valid assessment tool. The reliability coefficients range from .84 to .98 for the subtests and the IQ scales. The WASI has also been highly correlated with other overall cognitive ability measurements (Wechsler, 1999). The WASI can be administered as a four-subtest or two-subtest form; the two-subtest form was administered for this study. The two-subtest version of the WASI consists of the Vocabulary and Matrix Reasoning subtests and yields only the FSIQ score. Inter-correlations of FSIQ scores between the WASI two- and four-subtest forms are .94 for individuals 17-19 years old and .95 for individuals 20-24 years old. To

associated with Autism Spectrum Disorders (ASD) in individuals with normal intelligence. A Total AQ score of 32 or greater indicates a high probability of ASD. The Total AQ is comprised of 5 five subscales; Communication, Social, Imagination, Local Details, and Attention Switching. Test-retest reliability was shown to be good ($r=.7$) and self-report scores were shown to correlate with parent-reports of symptomatology. No participants were excluded from the current study based on the cutoff score of ≥ 32 .

Behavior Assessment System for Children, Second Edition, Self-Report - College (BASC-2, SRP-COL; Reynolds & Kamphaus, 2004). The Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004) is an assessment via self report of the participant's adaptive and problem behaviors in both the community and the home setting. The proposed study used the Self-Report – College (SRP-COL), which was appropriate for the age level of participants. The BASC-2, SRP-COL is a 185-item self-report questionnaire that assesses for a variety of behavioral difficulties. Domain scores are attained for: School Problems, Internalizing Problems, Inattention/Hyperactivity, Emotional Symptoms Index, and Personal Adjustment. The BASC-2 has strong psychometric properties to support its use. Research has shown good reliability. Internal consistency levels across normative samples range from .80 to .95 for composite scores and .80 to .87 for individual scales. These reliabilities are essentially the same in the clinical norm sample. The BASC-2 shows good test-retest reliability, with coefficients ranging from .80 to .90. Strong scale inter-correlations are indicative of the validity of the BASC-2. The BASC-2 correlates strongly with other similar measure of behavior. For example, the BASC-2 and Achenbach System of Empirically Based Assessment (ASEBA) have correlation coefficients ranging from .65 to .84. The composite and scale scores on the BASC-2 were used to assess the severity of any DSM-IV-TR symptoms and to determine

Table 1
Movie duration ranges in seconds.

Sentence	Minimum Duration	Maximum Duration
“Look in the box”	1.05	1.20
“Clouds are in the sky”	1.14	2.13
“It’s dark already”	1.10	1.50
“The dog is barking”	1.05	2.03
“The door is open”	1.08	1.23
“I didn’t expect you”	1.06	2.00
“It might happen”	1.05	1.29
“Put it down”	0.24	1.02
“It’s across the street”	1.07	2.02
“Turn off the television”	1.12	2.06

Screening Measures

The following cognitive, perceptual, and emotional measures were administered to all participants and the results determined eligibility for the current study. Analysis of these measures is beyond the scope of the current study and, therefore, was not included in current analyses. The number of participants excluded on the basis of performance is provided following psychometric description of each measure. Previous research suggests that individuals with psychopathology (i.e., anxiety, depression, psychosis) or a pervasive developmental disorder employ atypical face processing strategies. In addition, individuals with cognitive or perceptual disorders may employ atypical face processing strategies or be unable to complete an emotion decision task. As such, potentially atypical performance in this study may be better explained by perceptual or cognitive disorders than by tasks demands. Consequently, individuals who exceeded predetermined cutoffs for each of the following measures were excluded from the current study to reduce potential confounding variables.

The Autism-Spectrum Quotient (AQ; Autism Research Centre (ARC); Baron-Cohen et al. 2001). The AQ is a 50-item self-administered questionnaire that assesses for symptoms

right eye, nose, mouth; power $(1-\beta) = .99$). Participants ranged in age from 18.04-24.41 years ($M=19.81$, $SD=1.49$). The sample consisted of 42 females and 9 males. Participants were recruited from the undergraduate research pool at Georgia State University and received 3 hours course credit for their participation.

Experimental Stimuli

Dynamic audio-visual emotion (DAVE) stimuli were developed and validated at the Yale Child Study Center (Robins & Schultz, 2004). Movies were filmed and digitized with the help of a local production studio, using professional actors. Each stimulus consists of one actor delivering an emotionally ambiguous sentence, allowing for each emotion to be considered congruent with the semantic content. Actors delivered 10 different sentences in each of four emotional tones: angry, fearful, happy, and neutral. The complete set consists of 320 movies: 80 matching and 240 mismatching. Example sentences are: “The door is open” and “Look in the box.” To illustrate how the sentences are emotionally ambiguous, the sentence “Look in the box” could be construed as happy if one person is directing another to a birthday present, fearful if there is a snake in the box, angry if the glassware that someone ordered arrived broken, or neutral if one person were asking another if there was room to pack one more book. The digitized movies were then edited to create a set of clips with matching emotional presentations in face and voice (e.g., happy voice, happy face) and a set of clips with incongruent, mismatching emotional presentations in face and voice (e.g., happy voice, angry face). Digitization allowed for the synchronization of lip movement and sound. Each clip is approximately 2 seconds long. Movie duration ranges can be found in Table 1. Only movies between 1.00 and 2.13 seconds long were included in the current study.

would respond correctly more often than they would respond incorrectly (i.e., respond “yes” to congruent movies and “no” to incongruent movies). Because match/mismatch judgments would be relatively simple, it was expected that response times would be approximately equal during congruent and incongruent movie trials.

Hypothesis 8 – Task 3 Processing Approach. Because match/mismatch judgments would be relatively easy and automatic (similar to affective judgments for congruent movies), participants would again use a configural processing strategy more often than a featural strategy, when asked whether the emotion in the face matches the emotion in the voice. The configural processing approach would emerge during both the congruent and incongruent movies.

Hypothesis 9 – Task 3 Fixation Pattern. During match/mismatch judgments, participants would be overtly cued to integrate affective information from face (i.e., eyes) and voice (i.e., mouth); therefore, participants would fixate as often and for equal periods of time on the eye and mouth regions (which convey affective cues). This pattern of reliance on the eye and mouth regions would be the same for both the congruent and incongruent movies.

Method

Participants

Sixty participants were recruited for the current study and, in accordance with exclusion criteria, a final sample of 51 participants were included for analyses. A post-hoc power analysis was conducted to determine the power afforded by the sample size recruited for the present study using the program G*Power3 (Faul et al., 2007). According to the standards set by Cohen (1988), a sample size of 51 provided sufficient power to detect a medium effect using repeated measures ANOVA (within-subjects design) ($f=.25$, $\alpha=.05$) for two (e.g., core, periphery; power $(1-\beta) = .94$), three (e.g., eyes, nose, mouth; power $(1-\beta) = .93$), and four repetitions (e.g., left eye,

Hypothesis 5 – Task 2 Processing Approach. In contrast to the configural processing approach predicted to emerge during congruent movies, it was predicted that, when viewing incongruent movies, participants would use a featural processing strategy more often than a configural processing strategy. A featural processing strategy was defined as fixations being limited to the most salient region of the face (eyes) during more than 50% of the incongruent movies.

Hypothesis 6 – Task 2 Fixation Pattern. It was predicted that participants would again fixate for longer and more often on the eye region during the incongruent movies compared to other core regions, but that the difference between length and frequency of fixations for the eye and mouth regions would be greater during the incongruent movies than during the congruent movies. Because of the increase in relative complexity of incongruent movies as compared to congruent movies, individuals would increase fixations to the most salient facial feature for emotion judgments during incongruent movies, which was expected to be the eye region (Buchan et al., 2007; Lansing & McConkie, 1999; Rutherford & Towns, 2008). Specifically, it was expected that participants would increase the fixation time and number of fixations on the eye region, a primary source of affective information, during the incongruent movies as compared to congruent movies. Because the mouth is the feature most likely to be equated with the actor's voice, participants would decrease the fixation time and number of fixations on the mouth region during incongruent movies as compared to congruent movies.

Hypotheses pertaining to Task 3, which requires participants to make match/mismatch judgments in congruent and incongruent stimuli.

Hypothesis 7 – Task 3 Behavioral Responses and Reaction Time. When participants were asked whether or not the emotion portrayed in the face matches that expressed by the voice, they

a given condition. Exploratory analyses would also examine whether processing approach differs by emotion condition.

Hypothesis 3 – Task 1 Fixation Pattern. When viewing congruent movies, participants would fixate more often and for longer periods of time on core regions of the face (i.e., eyes, nose, mouth) than on peripheral regions (e.g., chin, forehead, cheeks). Because the eyes are expected to reveal the greatest amount of information about the affective state of another individual, the total fixation duration and fixation count will be greater in the eye region than in any other core region of the face. Because previous research utilizing the DAVE stimuli indicated different response patterns between the four emotions (angry, fearful, happy, neutral), additional analyses would explore potential differences in fixation patterns across emotions. In addition, if analyses indicated that participants showed a preference for one eye over the other, the left and right eyes would be analyzed separately for this and other analyses of fixation pattern.

Hypotheses pertaining to Task 2, which requires participants to label the emotion portrayed both in congruent and incongruent stimuli.

Hypothesis 4 – Task 2 Behavioral Responses and Reaction Time. When identifying the emotions portrayed in incongruent movies, participants would be significantly more likely to choose labels consistent with facial expressions than with vocal prosody (de Gelder & Vroomen, 2000; Hecht & Reiner, 2009; Massaro, 1998; Massaro & Egan, 1996). Response time to label emotions would be greater during the incongruent movie condition than during the congruent movie condition (Hietanen, Leppanen, Illi, & Surakka, 2004; Massaro & Egan, 1996; Pell, 2005).

The data collected in this study provide information as to the strategy employed (i.e., configural or featural) and the regions of the face (i.e., eyes, nose, mouth) that are sampled during a relatively more complex, and potentially more difficult, emotion perception task. In separate tasks, the participants were asked to identify the emotions portrayed and to determine whether the emotions were congruent or incongruent, with the hope of determining whether face processing strategies are instruction-dependent. The results of this study will serve as a comparison for future research extended to include one or more populations who have been shown to utilize atypical face processing strategies (e.g., Autism Spectrum Disorders, social phobia). Such a comparison would provide information as to where the emotion perception deficits arise or where the processing strategies are broken down. This knowledge would inform treatment intervention strategies that aim to improve social skills.

Hypotheses

Hypotheses pertaining to Task 1, which requires participants to label the emotion portrayed in congruent stimuli.

Hypothesis 1 – Task 1 Behavioral Responses. When identifying the emotions portrayed in congruent movies, it was predicted that participants would respond accurately more often than they would respond inaccurately.

Hypothesis 2 – Task 1 Processing Approach. Participants would use a configural processing approach more often than a featural approach when viewing congruent movies, scanning the entire face to elicit information about the configuration of facial features. A configural processing strategy was defined as fixations on both the eye region and the mouth region during more than 50% of the movies in one condition, whereas a featural strategy was defined as fixations limited to either the eye or the mouth region during the majority of movies in

revealed an opposite pattern; they decreased their foveal attention to the eye region, employing a hyper-scanning strategy and, as a result, significantly increasing their raw scanpath length (Horley et al., 2003, 2004). When presented with static photos of neutral faces, individuals with schizophrenia showed a reduced scanpath length and a trend toward fewer fixations than did typical controls (Manor et al., 1999). In another study, patients with schizophrenia were less able than controls to identify emotions portrayed by the eyes in isolation, but were able to identify emotions when presented with a whole face (Kington, Jones, Watt, Hopkin, & Williams, 2000).

The Current Study

In the current study, the researcher examined the gaze behavior of typically developing individuals when viewing emotional movie clips and making emotion judgments. Previous researchers have generally utilized static photographs in isolation when attempting to determine which regions of the face provide the most information for individuals under conditions requiring judgments of emotion. In studies designed to provide bimodal emotional information, printed or audio-taped recordings of sentences with emotional content or spoken in various tones of voice have been used (de Gelder & Vroomen, 2000; Grossman, Klin, Carter, & Volkmar, 2000; Hietanen, Leppanen, Illi, & Surakka, 2004; Lindner & Rosen, 2006; Massaro, 1998; Massaro & Egan, 1996; Pell, 2005). Dynamic stimuli were utilized in order to simulate the emotion decisions that are required during social interactions. The dynamic emotion stimuli are movie clips in which emotionally ambiguous sentences are presented by male and female actors who simultaneously portray an emotion via their facial expression and their tone of voice. Some of the movies have been digitally altered such that the emotion in the face does not match that in the voice. Gaze behavior during the congruent and incongruent emotion movie conditions was directly measured using eye-tracking technology.

Pelphrey and colleagues (2002) found that the visual scan-path of typical individuals resembled an upside-down triangle between the eyes, the nose, and the mouth, whereas the visual scan-path of individuals with ASD resembled a random pattern that focused on only one or two features surrounding the face (i.e., an ear, a region of the hairline, or the chin). When individuals with ASD did focus on a core feature, they selectively attended to the mouth (Pelphrey et al., 2002). Joseph and Tanaka (2003) found that, during emotion processing tasks, individuals with ASD depend more on information from the mouth region than from the eye region. Another group found that individuals with ASD did fixate on the eye region during the presentation of simple emotions at a rate similar to that of the group of typically developing individuals. As previously discussed, typically developing individuals increased fixations to the features, with a trend toward the eye region, when presented with complex emotions. In contrast, the individuals with ASD decreased the time spent looking at the facial features, with a trend toward decreased fixations devoted to the eyes (Rutherford & Towns, 2008).

One research group utilized dynamic stimuli to track the eye-scanning patterns of individuals with ASD. Klin and colleagues (2002) tracked the fixation time and patterns of individuals' eyes while watching an emotional movie segment. A typically developing individual responds directly to a look of surprise and horror in the individual's eyes and follows this nonverbal cue to the source. In contrast, an individual with ASD waits to follow the verbal cue and arrives at the wrong object (Klin et al., 2002).

Similar results have been found in individuals with social phobia and schizophrenia. As the emotional valence of another's face changes from neutral to positive (i.e., happy) to negative (i.e., sad, angry), a group of typically developing individuals increased their attention to the eye region, in both number of fixations and fixation duration. In contrast, those with social phobia

disproportionate sensitivity to face inversion; this results from the fact that these patients' feature-based strategy of face processing is unaffected by the disruption of the inverted face's configuration. In addition to individuals with autism and children, this feature-based strategy has been found in other populations. Individuals with prosopagnosia, a disorder characterized by the inability to recognize familiar or novel faces, have been shown not to exhibit an inversion effect (Behrmann et al., 2005). These patients perform the same, and perhaps better, on recognition of inverted faces relative to upright faces. In the study by Behrmann and colleagues (2005), individuals with congenital prosopagnosia were significantly less accurate than controls during an unfamiliar face recognition task and individuals with acquired prosopagnosia were significantly less accurate than both groups. As found previously, control individuals in this study demonstrated greater accuracy on recognition of upright than inverted faces. On the other hand, individuals with congenital prosopagnosia demonstrated no difference between recognition of upright and inverted faces, and individuals with acquired prosopagnosia showed a significant advantage for inverted faces. Children with Williams Syndrome have also failed to show an inversion effect (Deruelle et al., 1999); there was a decrease in recognition of faces, but this decrease was equal to the decrease in recognition of houses. In addition, the children with Williams Syndrome were shown to be less able than chronological and mental age-matched children to identify faces.

A number of researchers have shown that, in various psychiatric populations such as those with autism, social phobia, and schizophrenia, the failure to attend to the important information revealed by the eye region is related to social difficulties (Horley, Williams, Gonsalvez, & Gordon, 2003, 2004; Manor et al., 1999; Pelphrey et al., 2002). In a study in which participants were asked to identify the emotion portrayed by actors in static photographs,

of correctly identified emotions from static facial expression, dynamic facial expression, and prosody (neutral verbal content). The authors found that the AS group performed comparably to the control group in the content (neutral prosody) and the combined modalities (i.e., dynamic, prosody, and content) conditions, indicating the important role of verbal content as a mediator.

With regard to the difference in face perception between typically developing and children with autism, the latter have consistently demonstrated a bias for featural information and are less prone to use holistic information in the perception of faces, evidencing a whole-face recognition advantage that was more prominent during trials that relied on mouth-recognition (Behrmann et al., 2006; Joseph & Tanaka, 2003; Teunisse & de Gelder, 2003). As previously discussed, individuals with little experience with a class of objects, as opposed to those who have attained expertise, have been shown to rely more on the featural aspects of the face. Researchers that have compared the face processing strategies of children to adults have found that configural processing strategies develop more slowly than feature-based processing strategies, with a shift from analytic to holistic processing occurring with development (Mondloch, Le Grand, & Maurer, 2002; Schwarzer, Huber, and Dummler, 2005). Results indicated that children made significantly more errors, on a task requiring them to judge whether two faces were the same or different, than adults in a spacing task in which the space between eyes and nose differed (i.e., configural changes). In contrast, children were almost as accurate as adults in judgments of faces in which the external contour of the face changed and in judgments about faces with differing eye and nose shapes.

Recall that the inversion effect for faces is most commonly compared with the effects of inversion on other classes of stimuli. When comparing the magnitude of the inversion effect between two groups, researchers have found that some populations do not show such a

region, with the bilateral middle temporal gyri activating during configural processing (Lobmaier, Klaver, Loenneker, Martin, & Mast, 2008).

Atypical Face Processing and Emotion Perception

Much of what we have learned about face and emotion perception comes from populations that exhibit atypical processing strategies. Many researchers have attempted to understand deficits in emotion processing in specific psychiatric populations, particularly disorders in which such deficits are believed to impede social interaction. One such class of disorders is the Autism Spectrum Disorders (ASD). Weaknesses in bimodal integration have been found in individuals diagnosed with ASD. Gepner, de Gelder, and de Schonen (1996) presented participants with syllables on the auditory, visual, and audio-visual modalities and asked them to repeat exactly what the speaker said. They found that participants with autism were less efficient than matched controls on the visual-only and the auditory-visual conditions. Others have examined whether this deficit in integration extends to multimodal emotion processing. Grossman, Klin, Carter, and Volkmar (2000) studied individuals with Asperger's Syndrome by presenting photographs of actors portraying different emotions (e.g., happy, sad, fearful, anger, surprise) with matching, mismatching, and irrelevant printed words, requiring them to select the emotion word that described how the actor felt. The AS group was as accurate in responding to the matching condition, but significantly less accurate when the face was paired with a mismatching word. The AS group showed a clear bias toward the visual-verbal information, responding with the emotion that matched the printed word on the computer screen. Results from this and other studies suggest that individuals with ASD rely on verbal mediation as a compensatory strategy when they encounter complex emotion perception tasks. This idea was supported by Lindner and Rosen (2006), who found that individuals with AS had a lower number

degraded noise condition. Under the first condition, the greatest number of fixations was devoted to the eye region. In the second task, fixations were located significantly closer to the mouth than they were in the emotion condition; the authors concluded that the speech perception tasks drew participants' attention toward the mouth (Buchan et al., 2007).

Different scanning strategies may result from the interaction between lateralized brain function, tasks demands, and control of eye movement. Borod, Vingiano, and Cytryn (1988) found that providing emotion-laden instructions cause participants to look to the left side of the visual field. Individuals were asked to interpret verbal sentences (e.g., 'strike while the iron is hot') or imagine items (e.g., the Lincoln head of a penny) in the non-emotion condition. Individuals were asked to imagine experiences, presented in each of three modalities (i.e., visual, auditory, tactile), in the emotion condition. Examples of experiences that participants were asked to imagine include: seeing a brilliant sunset over the ocean, hearing a screeching siren approaching, and feeling a throbbing pain from a toothache. A researcher sat directly across from participants and rated the direction of lateral eye movements (LEMs). In contrast, others have found that participants exhibit a left visual field bias regardless of the type of instruction (Asthana & Mandal, 2001; Marcus, 2005). It has been postulated that the right-hemisphere role in emotion perception or emotion-related decision making becomes evident even when instructions are not emotionally laden (Asthana & Mandal, 2001). Other researchers have found support for lateralized influence on different face processing strategies. Some have argued that the left hemisphere, in general, is dominant for featural processing, whereas the right hemisphere is dominant for configural processing (Bourne, Vladeanu, & Hole, 2009). Featural processing has also been attributed more specifically to the left FG, left lingual gyrus, and the left parietal

more heavily on the external regions (e.g., the external face contour, forehead, lower face, chin), whereas configural-processors rely more on the internal regions (e.g., eyes, nose; Hobson et al., 1988; Joseph & Tanaka, 2003; Mondloch, Le Grand, & Maurer, 2002).

Records of eye movements have shown that foveal vision is reserved for elements containing essential information to the observer. In the case of face perception, most attention is paid to the eyes, nose, and mouth (Walker-Smith, Gale, & Findlay, 1977). Research with healthy individuals shows that participants first fixate on the eyes and mouth when they are provided with no task instructions (Bar-Haim, 2006; Groner, Walder, & Groner, 1984). Lansing and McConkie (1999) described the “Gaze Direction Assumption,” the theory that observers direct their gaze toward those parts of the stimulus from which visual information is being sought in order to carry out the task at hand. During emotion processing tasks, typical children and adults have a preference towards fixating on the eyes, mouth, and nose, regions argued to yield the greatest amount of information about the mental state of other individuals (Pelphrey et al., 2002; Walker-Smith, Gale, & Findlay, 1977). In the case of emotion perception, typically developing individuals looked significantly more towards the eye region than other regions when processing simple emotions (e.g., happy, surprise, angry, afraid). When presented with complex emotions (e.g., scheming, thoughtful, flirting), these individuals increased the time spent looking at the facial features, both eyes and mouth combined. The researchers found marginally significant effects, attributed to the weak statistical power, of increased time looking to the eyes, with decreased time to the mouth region (Rutherford & Towns, 2008). In another study, Buchan, Pare, and Munhall (2007) found that participant’s viewing strategies and regions of interest varied by task demands: one group was asked to identify the emotion portrayed (i.e., happy, neutral, or angry), while another group was required to identify all of the words spoken by the actor under a

Gaze Behavior

Noton and Stark (1971) stated that the most direct and real-time method of assessing processing strategies was to record visual scanpaths. The visual scanpath is defined as the pattern of eye movements that occurs when an individual processes a complex stimulus. Walker-Smith, Gale, and Findlay (1977) studied the eye movements of three participants to determine whether typical scanpaths emerged during a facial recognition task. Results showed that participants will normally adopt a regular and consistent scanning strategy when viewing faces for encoding and during recognition. Research examining gaze behavior and eye movements during viewing of faces regularly indicates a proportionally greater sampling of the internal region compared with the external region and point of regard primarily to the core features (Luria & Strauss, 1978; Macworth & Bruner, 1970; Manor et al., 1999; Mertens, Sigmeund, & Grusser, 1993; Stacey et al., 2005).

Schwarzer, Huber, and Dummler (2005) explored the relationship between gaze behavior, specifically the information used at the stage of visual encoding, and different methods of face processing. Findings indicated that individuals utilizing a holistic and configural approach to face processing directed a substantially larger proportion of fixations and time to the eyes and the nose. On the other hand, those who processed the face analytically, employing a featural processing strategy, directed a greater number of fixations and time to one specific feature that they later used for processing and these fixations were more distributed across facial features (Schwarzer, Huber, and Dummler (2005). Other researchers have reported similar findings. Individuals who rely on featural processing, such as children, individuals with autism, or prosopagnosic patients, attend to different facial regions when recognizing faces than do those who rely on configural processing. Some researchers have argued that featural-processors rely

that class of stimuli (Gauthier et al., 1998; Gauthier & Tarr 1997). In addition to human faces, the ability to achieve expertise has been demonstrated with other stimulus classes such as dogs and birds (Diamond & Carey, 1986; Tanaka & Taylor, 1991).

The question of whether the superior face processing abilities of humans result from the unique specialization of cortical areas (e.g., FFA) or from the extensive amount of experience we have with this class of stimuli, qualifying us as experts, is yet to be answered. Researchers continue to explore these two possibilities, with both theories subject to sound methods and supported by convincing data. Recent researchers have utilized neuroimaging and physiological measures in an attempt to further our understanding of the neural underpinnings of face and emotion perception. Related to the previously discussed face processing strategies, Itier, Alain, Sedore, and McIntosh (2007) found that disruption of the facial configuration, by inversion or contrast reversal, results in the increased amplitude and latency of a face-sensitive marker: the N170 ERP. Previous work by this group indicated that three neural regions are active during the N170: the superior temporal sulcus (i.e., STS region), the fusiform gyrus, and an occipital area (Itier et al., 2007). The authors postulate that the STS region is the primary region associated with the N170 and that face- and eye-selective neurons exist here. They argued that, in upright faces, the preserved configuration activates face-selective neurons. The change in the negative N170 ERP occurs when additional eye-selective neurons are recruited to process the eye region because of the disrupted facial configuration (Itier, Alain, Sedore, and McIntosh, 2007). The STS region, along with the superior temporal gyrus (i.e., STG), has also been implicated in the integration of visual and auditory perceptual cues and in emotion perception (Robins, Hunyadi, & Schultz, 2009).

stimuli can be classified at the individual Greeble level, but also grouped into family and gender, thus essentially equating them to human stimuli. In the development of the Greeble stimuli, the characteristics of face recognition were taken into account. Like human faces, the Greeble stimulus set consists of individuals with features that can be studied in isolation or in configuration. The characteristics of face recognition include: similar features are organized in a similar configuration, faces are typically recognized at the exemplar-specific level, and humans are experts at discriminating human faces (Gauthier & Tarr, 1997).

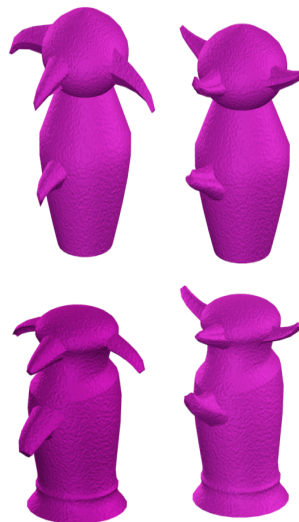


Figure 1. Four Greebles: two female on the left and two male on the right, with the upper and lower stimuli belonging to separate families.

With regard to the second characteristic, Gauthier and Tarr (1997) postulated that most objects are recognized at the basic level (e.g., boy), whereas faces are recognized at the subordinate, or individual, level (e.g., Ryan). Expertise with faces results in individuals' judgments of face identity, a subordinate level decision, being as quick as judgments of gender or race, a category or basic level decision. Expertise with a class of objects is said to be attained when the individual achieves this qualitative shift in processing: the stimulus is recognized as quickly at the subordinate level as it is at the basic level and the FFA is selectively activated for

and to make certain that activation was to the face in general, as opposed to lower-level visual features (i.e., specific features that are common to all faces; Kanwisher & Yovel, 2006; Tong et al. 2000). During studies of cortical activation in response to viewing face stimuli, some researchers have also found activation in the middle temporal gyrus or superior temporal sulcus (STS), and an area of the lateral occipital cortex that came to be known as the “occipital face area” (OFA; Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott, & Chun, 1997; Kanwisher, & Yovel, 2006). The roles of these three areas, FFA, STS, and OFA, were compared in various tasks and it was determined that each extracts different types of information from face stimuli. The FFA and the OFA are implicated in tasks requiring individuals to distinguish between individual faces, whereas the posterior STS (pSTS) is recruited during tasks pulling for the processing and recognition of emotional expression and eye gaze (Haxby, Hoffman, & Gobbini, 2000; Kanwisher, & Yovel, 2006). Because of the robust and consistent findings of FFA activation across studies, this area continues to be the primary cortical region studied in relation to face perception.

Others have argued that this specialization is not unique to faces, but that participants possess superior processing abilities for classes of objects for which expertise has been developed. Following the argument of Diamond and Carey (1986), Gauthier and Tarr (1997) explored the role of expertise in face recognition by using novel, non-face stimuli in a recognition task (“Greebles”). Two brain regions, previously discussed, are selectively and significantly more active as expertise (e.g., with Greebles) is acquired: the FFA and the OFA (Bukach, Gauthier, & Tarr, 2006; Gauthier et al., 1999; Tarr & Gauthier, 2000). Greebles are three-dimensional figures that serve as control stimuli for faces. These images, courtesy of Michael J. Tarr, Brown University, <http://www.tarrlab.org/>, can be viewed in Figure 1. These

expertise included evidence of holistic processing and configural processing. Parts were better recognized in the studied configuration than when presented in isolation, indicating that they were encoded holistically. Furthermore, extensively trained viewers (i.e., experts) were more sensitive to configuration change than were novice viewers, indicating that expert viewers encoded the stimuli configurally (Gauthier & Tarr, 1997).

Neural Mechanisms

Whether humans show superior face recognition and processing, as compared to other classes of objects, because of a special or unique process or because they have obtained expertise continues to be a central theoretical issue in the face processing literature. Researchers have begun to explore this question through imaging studies. Similar to the argument proposed by Yin (1969, 1970), Kanwisher, McDermott, and Chun (1997) argued that there are specific regions of the cortex that are specialized for face perception. They conducted a series of experiments and results indicated that a region on the lateral side of the mid-fusiform gyrus responded significantly more strongly during the passive viewing of a human face than other objects (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Moscovitch, 2000). This region was identified as the “fusiform face area” (FFA) and researchers identified this as a specific region of interest (ROI) to examine in subsequent analyses. They found that this face-specific ROI was activated more strongly when viewing intact two-tone faces than scrambled same-tone faces, the average percentage signal increase, from fixation baseline, of the face ROI was greater for faces than for houses, and that the face ROI activated more strongly when passively viewing and matching three-quarter faces than hands (Kanwisher, McDermott, & Chun, 1997).

The findings that the FFA was specifically activated for faces were later extended and replicated to include a wide variety of face and face-like stimuli (i.e., cartoon faces, cat faces)

importance when considering the superior recognition or processing of faces as compared to other classes of objects and a number of theories have been proposed that attempt to explain how this information is represented. In addition to the theories of relational or configural information, others have proposed that gestalt form of the face is represented in parallel processing or through interactive coding (Bradshaw & Wallace, 1971; Diamond & Carey, 1986; Rhodes, 1988; Rhodes, Brake, & Atkinson, 1993; Sergent, 1984). Farah, Wilson, Drain, and Tanaka (1998) have attempted to explain how all of the parts of a face are processed, postulating that face recognition does not involve, or does so only to a lesser extent, the explicit representations of the eyes, nose, and the mouth. They believe that faces are instead recognized as undifferentiated wholes. In support of this theory, Farah and colleagues (1998) found that individuals perceived upright faces in a holistic manner, significantly more so than inverted faces, houses, or words; participants had significant difficulty differentiating and comparing individual parts of the upright face independent of the whole face.

Others have found a similar whole-face advantage for upright faces. Homa, Haver, and Schwartz (1976) found that individuals recognized features more accurately when presented in the context of a natural face than if presented amongst scramble features. Typically developing children recognized eyes and mouths in the context of the whole face, with eyes being recognized more efficiently than the mouth (Joseph & Tanaka, 2003). If these component features were isolated from the whole context of the face, recognition was more disrupted for upright faces than inverted or scrambled faces. This isolation was also more disrupted than the isolation of component features of houses (Tanaka & Farah, 1993). Gauthier and Tarr (1997) also examined recognition along this isolation-configuration continuum with their novel stimuli. In addition to the subordinate level shift previously described, two other behavioral markers of

defined by Diamond and Carey (1986), first-order relational features refer to the relationship between parts of a stimulus that are relatively unconstrained (i.e., the trees in a landscape). When you have a class of stimuli that are homogenous in nature (i.e., human faces), the first-order relational features are relatively constrained (i.e., eyes are located horizontally above the nose, which is above the mouth; Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993). With stimulus classes that possess fixed first-order relational features, another relationship becomes crucial to the individuation of class members. The distinctive relations between the fixed features of a stimulus (e.g., the distance between a specific individual's eyes) are referred to as the second-order relational features; this unique configuration is what is used to individuate faces (Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993).

There is a growing body of literature focused on contrasting the configural versus featural processing of objects, human faces in particular. A common assumption is that both information about isolated features (i.e., featural properties) and the relations among them (i.e., configural properties) are used as a means of encoding upright faces. The configural information, however, cannot be extracted from inverted faces and therefore results in less successful recognition (Valentine, 1988). Diamond and Carey (1986) argued that the large inversion effect that results from expertise and experience is due to individuals' reliance on the configural properties of stimuli (e.g., faces). Researchers have shown that the coding of second-order relational features is more affected by inversion than the coding of first-order relations or of isolated features (Rhodes, Brake, & Atkinson, 1993). Rhodes (1988) concluded that second-order features are critical to face perception and thus it is, at least in part, configural by nature.

In addition to the second-order relations, the properties of a stimulus configuration include the holistic, or gestalt, form. The overall structure or gestalt of human faces is of critical

class of stimuli. Second-order relational features, to be discussed in detail in the next section, essentially are to the spatial configuration of the individual features of a stimulus. With regard to the criterion of expertise, Diamond and Carey (1986) found that the inversion effect emerges when perceivers are as experienced with the class of stimuli presented for recognition as they are with human faces. They found that the recognition of dogs is as sensitive to the inversion effect as is the recognition of faces, provided that the perceivers are highly experienced with both dogs, such as dog breeders, and faces. This argument was in contrast to an earlier theory proposed by Yin (1969, 1970). Yin (1969, 1970) found that memory recognition for human faces is more vulnerable to stimulus inversion when compared to other classes of familiar objects (e.g., houses, airplanes, stick figures of people in motion) and argued that the inversion effect occurs because neural specialization has evolved such that humans possess superior recognition processing strategies for human faces. The first criterion proposed by Diamond and Carey (1986), regarding stimulus-class characteristics, provides insight into the underlying processes of facial and emotion processing and will be discussed in greater detail.

Distinguishing between first-order and second-order features is important to understanding the differential effects of inversion on recognition. Rhodes (1988) defined first-order features as the appearances of individual parts of the face, such as those labeled the eye, nose, and mouth. Second-order features were those said to have configural properties, including the spatial relations among first-order features, the position of first-order features, and information about the shape of the face (Rhodes, 1988). Rhodes' initial hypothesis of the mechanisms of face representation utilized similar terminology as that proposed by Diamond and Carey (1986), but the definitions appeared somewhat different. The distinction between the two hypotheses became less clear in a later paper written by Rhodes, Brake, and Atkinson (1993). As

insight into typical and atypical bimodal emotion perception. Research indicates that a global or configural representation of faces was related to a greater recognition of facial emotion (Gross, 2005; Smith & Scott, 1998; White, 2000). A global or configural representation refers to one in which the perceiver utilizes information about the position of facial features, or configuration of the features, to a greater degree than information about the individual features of the face. To identify a particular face or facial expression, typically developing individuals generally rely on the spatial configuration of core facial features (i.e., eyes, nose, mouth), utilizing a holistic processing approach, as opposed to relying on individual features, associated with the processing of objects (Pelphrey et al., 2002). Exploring these various approaches to face processing (i.e., configural, featural, holistic) in greater detail is important to furthering our knowledge of how and what information is extracted from the face during emotion processing tasks.

Face Processing / Perception

There is a large body of literature that is dedicated to the exploration of face processing strategies, particularly configural and featural strategies. Diamond and Carey (1986) are some of the researchers at the forefront of this literature and have attempted to explain the inversion effect, a manipulation frequently utilized in standard recognition memory paradigms. The inversion effect is essentially a difference in performance on recognition tasks between upright and inverted photographs of faces; the inversion effect occurs when faces that are presented upside-down during encoding and recognition are more poorly recognized than faces presented upright.

Diamond and Carey (1986) argued that the inversion effect emerges when two criteria emerge: (1) members of a class to be encoded share a configuration and can be distinguished from one another based on second-order relational features and (2) perceivers are experts with a

Similar to the manner in which auditory and visual syllables are integrated, in the manner illustrated by the McGurk Illusion, conflicting emotional information from the face and voice are integrated. An example that illustrates a unique emotional percept is that of sarcasm. An individual may state, “I would love to have lunch with you” with a smile on his face, but one may perceive negative affect in his tone of voice. If an individual fails to attend to the prosody and attends only to the facial expression of happy, he or she will misconstrue the intent of the person who delivered the comment.

Emotion identification decisions are slower and less accurate when facial expression and prosody are incongruent (Hietanen et al., 2004; Massaro & Egan, 1996; Pell, 2005). A number of researchers have demonstrated that typically developing individuals will show a bias toward facial information when incongruent information is presented in another modality (de Gelder & Vroomen, 2000; Massaro, 1998; Massaro & Egan, 1996). When participants are instructed to attend to both the visual and auditory modalities, the emotional attributes of the face had a more significant effect on responses than did the prosody (Massaro & Egan, 1996). These findings were replicated by Massaro (1998), showing that the auditory information had no appreciable influence on the recognition of the facial affect, with the exception that a marginally significant effect was found for neutral faces. On the other hand, Massaro (1998) found that the visual information had a distinct influence on the recognition of emotion by voice. Another research group found that unattended facial affect influences the perception of prosodic information, both when the face onset was delayed from the spoken sentence and when they were presented concurrently (de Gelder & Vroomen, 2000; Vroomen, Driver, & de Gelder, 2001).

Given that information from the visual modality appears to have a greater influence on audio-visual integration, exploration of facial processing strategies is bound to offer critical

(Hietanen, Leppanen, Illi, & Surakka, 2004). When congruent visual and auditory information is presented concurrently, this integration occurs, as measured by event related brain potentials (ERPs), as early as 110 ms after stimulus onset (Pourtois et al., 2000).

In order to understand how people integrate auditory and visual cues, it is helpful to look at the broader, more studied field of perceptual integration. The integration of mismatching sounds and visual input is important to consider when investigating emotion processing. Individuals are required to integrate information from multiple channels that are sometimes in conflict with one another. The observer is responsible for discriminating and making sense of the information when in conflict. Evidence for this early audio-visual integration can be found in research on the McGurk Illusion. This phenomenon demonstrates how multimodal information is combined to create a unique percept (McGurk & MacDonald, 1976). When an auditory syllable (i.e., /iki/) is presented simultaneously with a visual syllable (i.e., /ipi/), a third syllable is perceived (i.e., /ipki/ or /ikpi/). Another example of how the two initial syllables can be integrated is when an auditory syllable (i.e., /ipi/) and a visual syllable (i.e., /ikli/) are fused together to create another syllable (i.e., /iti/ or /idi/). Some researchers have found that one sensory modality has a greater influence on perceptual outcome than others. When presented simultaneously with visual and auditory stimuli and asked to respond by indicating whether they detected a visual stimulus, an auditory stimulus, or both, researchers found that participants made a significant number of errors. Specifically, participants were prone to report visual only during trials in which both visual and auditory cues were present, termed the visual dominance effect (Hecht & Reiner, 2008). In another study, this visual dominance effect emerged both when stimuli were semantically congruent and when they were semantically incongruent (Koppen, Alsius, & Spence, 2008).

Introduction

The ability to produce and recognize facial expressions is an important factor in successful social encounters and interpersonal relationships. Facial expressions of others reveal a lot of information about their state of mind, their desires, and their intentions. Affective perception occurs quickly and requires the seemingly automatic integration of facial expression with additional sources of information. These other sources may include, but are not limited to, verbal information, affective prosody, affective memory, and physiological responses.

Researchers continue to explore the manner in which humans derive affective information from the facial expressions of others, but the process is not yet completely understood. In order to fully understand the perception of visual affective information, it is important to consider the underlying processes. In accordance with this, research on emotion perception will be reviewed, taking into consideration what is known about the perception of visual, auditory, and audio-visual information. In addition, a broader perceptual literature provides a useful foundation from which predictions can be made about the strategies utilized during emotion processing. Research on facial processing strategies and gaze behavior during a variety of tasks provides a foundation for understanding the manner in which essential information is gathered about the affective state of others. Although beyond the scope of the current study, it is also critical to explore the neural mechanisms that underlying such processes, as this may inform future studies.

Emotion Perception

In everyday social situations, individuals are required to integrate information from both the visual (i.e., facial expression) and the auditory (i.e., tone of voice) modalities. Evidence shows that the integration of facial and vocal information occurs early on in emotion perception, during the perceptual level of processing and not later during the response selection stage

<i>Figure 23.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 3.	68
<i>Figure 24.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 3.	69
<i>Figure 25.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 3.	70
<i>Figure 26.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 3.	71

<i>Figure 12.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 2.	57
<i>Figure 13.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 2.	58
<i>Figure 14.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent happy movies in Task 2.	59
<i>Figure 15.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 2.	60
<i>Figure 16.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent neutral movies in Task 2.	60
<i>Figure 17.</i> Total fixation duration to the two areas of interest during congruent and incongruent movies of Task 3.	64
<i>Figure 18.</i> Total fixation count to the two areas of interest during congruent and incongruent movies of Task 3.	64
<i>Figure 19.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 3.	65
<i>Figure 20.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 3.	66
<i>Figure 21.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 3.	67
<i>Figure 22.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 3.	68

LIST OF FIGURES

<i>Figure 1.</i> Four Greebles: two female on left and two male on the right, with the upper and lower stimuli belonging to separate families.	10
<i>Figure 2.</i> Schematic of ASL 6000 set-up and eye-tracking data collection.	32
<i>Figure 3.</i> Sample of male and female stimuli with defined facial regions.	34
<i>Figure 4.</i> Total fixation count to the core and peripheral regions during congruent movies of Task 1.	44
<i>Figure 5.</i> Total fixation duration to the four core regions of the face during congruent movies of Task 1.	47
<i>Figure 6.</i> Total fixation count to the four core regions of the face during congruent movies of Task 1.	49
<i>Figure 7.</i> Total fixation duration to the three areas of interest during congruent and incongruent movies of Task 2.	53
<i>Figure 8.</i> Total fixation count to the three areas of interest during congruent and incongruent movies of Task 2.	53
<i>Figure 9.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 2.	55
<i>Figure 10.</i> Total fixation count for the left eye, right eye, and mouth when viewing congruent and incongruent angry movies in Task 2.	55
<i>Figure 11.</i> Total fixation duration for the left eye, right eye, and mouth when viewing congruent and incongruent fearful movies in Task 2.	56

Table 13 <i>Descriptive statistics of fixation duration (sec) and count during fearful movies of Task 2</i>	56
Table 14 <i>Descriptive statistics of fixation duration (sec) and count during happy movies of Task 2</i>	58
Table 15 <i>Descriptive statistics of fixation duration (sec) and count during neutral movies of Task 2</i>	59
Table 16 <i>Proportion of participants utilizing configural or featural processing strategies of Task 3</i>	62
Table 17 <i>Descriptive statistics of fixation duration (sec) and fixation count during congruent and incongruent movies of Task 3</i>	63
Table 18 <i>Descriptive statistics of fixation duration (sec) and count during angry movies of Task 3</i>	65
Table 19 <i>Descriptive statistics of fixation duration (sec) and count during fearful movies of Task 3</i>	67
Table 20 <i>Descriptive statistics of fixation duration (sec) and count during happy movies of Task 3</i>	68
Table 21 <i>Descriptive statistics of fixation duration (sec) and count during neutral movies of Task 3</i>	70

LIST OF TABLES

Table 1 <i>Movie duration ranges in seconds</i>	25
Table 2 <i>Contrast between fixation duration for the left and right eye</i>	39
Table 3 <i>Proportion of participants utilizing configural or featural processing strategies of Task 1</i>	41
Table 4 <i>Descriptive statistics for fixation data during congruent movies of Task 1</i>	43
Table 5 <i>Descriptive statistics of fixation duration (sec) and fixation count during congruent movies of Task 1</i>	45
Table 6 <i>Descriptive statistics of fixation duration (sec) for congruent emotion movies of Task 1</i>	46
Table 7 <i>Within-subjects contrasts for the Emotion x Region interaction for fixation duration (sec) to the core regions in Task 1</i>	47
Table 8 <i>Descriptive statistics of fixation count during congruent movies of Task 1</i>	48
Table 9 <i>Within-subjects contrasts for the Emotion x Region interaction for fixation count to the core regions in Task 1</i>	49
Table 10 <i>Proportion of participants utilizing configural or featural processing strategies of Task 2</i>	51
Table 11 <i>Descriptive statistics of fixation duration (sec) and fixation count during congruent and incongruent movies of Task 2</i>	53
Table 12 <i>Descriptive statistics of fixation duration (sec) and count during angry movies of Task 2</i>	54

Hypothesis 6 – Task 2 Fixation Pattern	51
Hypothesis 7 – Task 3 Behavioral Responses and Reaction Time	61
Hypothesis 8 – Task 3 Processing Approach	61
Hypothesis 9 – Task 3 Fixation Pattern	62
Discussion	71
Behavioral Responses and Response Time	72
Processing Approach	75
Lateralization of Emotion Perception	78
Fixation Patterns: Congruent and Incongruent Movies	84
Fixation Patterns: The Effect of Stimulus Affect on Emotion Judgments	87
Fixation Patterns: The Effect of Emotion on Match/Mismatch Judgments	92
Conclusions	93
References	97
Appendix 1	110
Appendix 2	111

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
Introduction	1
Emotion Perception	1
Face Processing / Perception	4
Neural Mechanisms	8
Gaze Behavior	12
Atypical Face Processing and Emotion Perception	15
The Current Study	19
Hypotheses	20
Method	23
Participants	23
Experimental Stimuli	24
Screening Measures	25
Apparatus	30
Procedure	32
Results	38
Hypothesis 1 – Task 1 Behavioral Responses	40
Hypothesis 2 – Task 1 Processing Approach	40
Hypothesis 3 – Task 1 Fixation Pattern	41
Hypothesis 4 – Task 2 Behavioral Responses and Reaction Time	50
Hypothesis 5 – Task 2 Processing Approach	50

GAZE FIXATION DURING THE PERCEPTION OF VISUAL AND AUDITORY
AFFECTIVE CUES

by

SUSAN M. MCMANUS

Committee Chair: Diana L. Robins, Ph.D.

Committee: Erin B. Tone, Ph.D.

David A. Washburn, Ph.D.

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2010

Copyright by
Susan M. McManus
2010

GAZE FIXATION DURING THE PERCEPTION OF VISUAL AND AUDITORY
AFFECTIVE CUES

by

SUSAN M. MCMANUS

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2010

tasks, with a significant preference for the eyes. Evidence of hemispheric lateralization was indicated by preferential fixation to the left (happy, angry) or right eye (fearful). Fixation patterns differed according to the facially expressed emotion, with the pattern that emerged during fearful movies supporting the significance of automatic threat processing. Finally, fixation pattern during the perception of incongruent movies varied according to task instructions.

INDEX WORDS: Emotion perception, Eye-tracking, Gaze behavior, Face processing

GAZE FIXATION DURING THE PERCEPTION OF VISUAL AND AUDITORY AFFECTIVE CUES

by

SUSAN M. MCMANUS

Under the Direction of Diana L. Robins, Ph.D.

ABSTRACT

The accurate integration of audio-visual emotion cues is critical for social interactions and requires efficient processing of facial cues. Gaze behavior of typically developing young adults was measured via eye-tracking during the perception of dynamic audio-visual emotion (DAVE) stimuli. Participants were able to identify basic emotions (angry, fearful, happy, neutral) and determine the congruence of facial expression and prosody. Perception of incongruent videos resulted in increased reaction times and emotion identification consistent with the facial expression. Participants consistently demonstrated a featural processing approach across all