Published in final edited form as:

Psychon Bull Rev. 2014 December ; 21(6): 1537–1543. doi:10.3758/s13423-014-0629-y.

Detecting Personal Familiarity Depends on Static Frames in “Thin Slices” of Behavior

Alyson Saville and Benjamin Balas

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University

Abstract

Brief glimpses of nonverbal behavior (or ‘thin slices’) offer ample visual information to make reliable judgments about individuals. Previous work has largely focused on personality characteristics and traits of the individual; however the nature of dyadic relationships (strangers, lovers, or friends) can also be determined (Ambady & Gray, 2002). Judgments from thin slices are known to be accurate, but the motion features supporting accurate performance are unknown. We explored whether personal familiarity was detectable within the context of ‘thin slices’ of genuine interaction and, the invariant properties of thin-slice recognition. In two experiments, participants sequentially viewed two 6-s silent videos on each trial of an individual interacting with an unfamiliar partner; the other depicted the same person interacting with a personally-familiar partner. All sequences were cropped so that only the target individual was visible. In Experiment 1, participants either viewed the original sequences, reversed sequences, a static-image “slideshow” of the sequence, or a static-image slideshow with blank frames separating each image. In Experiment 2, all participants viewed the original sequences and either clips played at double-speed or half-speed. Participants’ performance was above chance in the forward and reverse conditions, but was significantly better in both the static-image slideshow conditions. When task speed was manipulated, we found a larger performance cost for fast videos compared to slow videos. Detecting personal familiarity via spontaneous natural gesture depends on information in static images more than face or body movement. While static images are typically less important for recognizing nonverbal behavior, we argue they may be valuable for making familiarity judgments from thin slices.

Keywords

Social Cognition; Thin Slice Perception; Motion Processing

Introduction

The capacity to make inferences about others is a valuable tool in human interaction and can be done quickly and accurately. Glimpses of nonverbal behavior (‘thin slices’) offer ample information to make rapid and reliable judgments about individuals including: teacher

effectiveness (Ambady & Rosenthal, 1993), sales successfulness (Ambady, Krabbenhoft, & Hogan, 2006), racial attitudes (Richeson & Shelton, 2005) and sexual orientation (Rule & Ambady, 2008). Observers draw conclusions based on brief observations and are often very accurate when judged against objective measures of performance and subjective measures of the same traits. Suppression of nonverbal behavior is difficult and thus gesture, posture, and expression offer valuable information regarding the internal disposition of an individual (DePaulo, 1992).

Much of the thin-slice research has focused on individual personality traits; the ability to apprehend dyadic interactions has been less explored. Relative status and kinship can be deduced from brief video clips (Costanzo & Archer, 1989), and interaction paradigms have also been used to understand how people in relationships identify with one another (Ickes, 2001). In a study done by Ambady & Gray (2002), observers viewed silent interactions between opposite-sex dyads and were asked to determine the nature of the relationship (strangers, lovers, or friends). Participants could identify the relationship status of the pair at above chance levels suggesting the nature of dyadic relationships can be successfully determined. Thin-slice perception thus extends beyond individual qualities and includes the ability to make inferences about the social characteristics of interactions between individuals. Presently, we chose to explore how observers use different visual cues to make such judgments. Specifically, we wished to determine what visual information is used for thin-slice judgments of personal familiarity between individuals.

We elected to study mutual familiarity between individuals because of its ecological relevance to everyday social interaction. Specifically, we chose to use candid, natural interactions so that the present research would be representative of real-world exchanges amongst individuals. In certain instances observers are able to determine familiarity from interactions displayed in silent films - the exhibition of different behaviors in our interactions with familiar and unfamiliar individuals via nonverbal channels offers important signals in the recognition of others (Feyereisen, 1994), though previous results suggest that perhaps only limited abilities to recognize mutual familiarity are evident. Early results suggested that the evaluation of mutual familiarity in a dyadic interaction given only a short video depicting one of the actors led to non-random response distributions (Benjamin & Creider, 1975), but offered little quantitative evidence that performance was veridical. Abramovitch (1977) asked children to watch film segments of their mother interacting with a familiar and unfamiliar partner and an unknown person engaging in similar interactions. Observers' judgments were only greater than chance when their own mother was a target. Detecting personal familiarity is thus non-trivial and may rely on a range of behavioral cues including inter-personal distance, mutual contact, etc., which differ in unconstrained settings (Feyereisen, 1994).

While a variety of thin-slice judgments can be made reliably, it remains largely unknown what visual information observers use to accomplish these tasks. Our goal was thus to examine the impact of natural motion and static appearance on thin-slice perception by manipulating the availability of natural motion and static frames in a mutual familiarity recognition task. Video clips offer information about dynamic behaviors that include gestures and body movements, whereas static appearance conveys information regarding

expressions and appearance with minimal movement. Movement facilitates recognition performance (Schiff, Banka, & de Bordes Galdi, 1986), and thin-slice judgments such that accuracy is greater when dynamic cues are employed as opposed to only static information (Ambady, Conner, & Hallahan, 1999) and also is greater when temporal order is preserved (Balas, Kanwisher, & Saxe, 2012). We sought to address the specific features of motion that are utilized in making successful thin-slice determinations within the context of personal familiarity. In two experiments, we manipulated the playback direction, rate, and smoothness of stimuli to determine how static frames and coherent motion sequences contribute to our ability to accurately predict behavior from thin slices. In contrast to previous reports, we find evidence in both tasks that static frames may carry particularly useful information for thin-slice perception. We discuss our results in the context of person recognition in general, and suggest processes that may support thin-slice recognition in particular.

Experiment 1

In our first experiment, we recorded multiple dyads engaged in a short interactive task and asked independent observers to categorize individuals' non-verbal behaviors according to their perceived familiarity with an unseen partner. We manipulated the motion content of these stimuli by playing them forwards or backwards, or in two "slideshow" conditions with minimal motion cues.

Method

Participants

Sixty-four young adults (10F/6M in the forward video condition, 12F/4M in the reverse video condition, 10F/6M in the slideshow condition, and 12F/4M in the blank slideshow condition) between the ages of 18-34 years participated in Experiment 1. Participants were students at North Dakota State University who received course credit or pay for participating.

Stimuli

We created thin-slice stimuli by videotaping 16 young adults interacting while playing Legos with someone familiar (e.g., a good friend or significant other) and unfamiliar to them. Each pair was instructed to construct complementary Lego buildings during a 10 minute period while seated across from each other at a small table. A video camera was stationed towards the back wall of an enclosed chamber to capture a side view of participants' interactions. Recording began immediately after instructions were given. Participants wore grey t-shirts and recording conditions were matched across dyads.

Our full set of videos comprised 16 unique pairs of individuals, including interactions between same-gender (5 male/male and 7 female/female) and opposite-gender partners (4 female/male). We extracted nine 6-second clips from each unique interaction (a total of 144 clips) which included obvious social behavior, specifically eye contact and speech. This initial selection was intended to exclude segments in which participants did not move.

Pilot testing with the full stimulus set demonstrated that participants were not above chance at categorizing videos by familiarity, making it impossible to assess the impact of our proposed manipulations. We therefore used the Amazon Mechanical Turk to obtain familiarity ratings for our clips on a 1-7 Likert scale to select a subset of video clips that maximally differed according to perceived familiarity. We identified clips of the same individuals in highly-rated “Familiar” videos ($M = 5.21$, $SE = 0.07$) and low-rated “Unfamiliar” videos ($M = 4.37$, $SE = 0.06$). We chose a subset of 60 video clips from our original database of 144 clips based on this criterion. Critically, our use of pre-selection means that above-chance categorization for unaltered videos is not an interesting result in itself. The impact of stimulus manipulations is however, the focus of this study, and depends critically on establishing above-chance baseline performance.

We ran all experiments using the Matlab Psychophysics Toolbox (Brainard, 1997) on a 13-inch MacBook (1280×800 resolution, 60Hz refresh-rate) with a 2.4 GHz Intel Core 2 Duo processor. Stimuli consisted of 6-second clips, cropped in half (540×480 pixels) so only one individual was present in the video. The experiment consisted of thirty trials presented in a random order and each trial was comprised of two matched silent 6-s videos of the same person (one familiar clip and one unfamiliar clip) presented sequentially in a random order (Figure 1). Viewing distance was approximately at 50 cm.

Procedure

We informed participants in all conditions they would see several pairs of short videos depicting people playing with an unseen partner and that within each pair, one video would depict a “Familiar” interaction and the other would depict an “Unfamiliar” interaction. Participants were asked to decide in which video (1 or 2) they believed the “Familiar” interaction occurred. Response time was unlimited and responses were made with the keyboard. In the “Reversed” condition, videos were played backwards. In the “Slideshow” condition, the stimuli consisted of 8 frames presented for 800ms each that were sampled uniformly from the original clips (Figure 2). In the “Blank Slideshow” condition participants viewed the same 8 frames as in the “Slideshow” condition, with a blank screen presented for 133ms between each frame to minimize apparent motion between static frames. We included this latter condition because apparent motion is known to be a necessary and sufficient cue for interpreting facial expressions (Ambadar, Schooler & Cohn, 2005).

Results

We analyzed participants’ accuracy in the forward, reverse, slideshow and blank-slideshow conditions using a 2×4 between-subjects analysis of variance (ANOVA) with sex (female or male) and video condition as between-subjects factors. This revealed a main effect of task, $F(3,56) = 3.30$, $p = .027$, partial eta-squared = .150, but sex, $F(1, 56) = .007$, $p = .933$, partial eta-squared = .001, and the sex by task interaction, $F(3, 56) = 2.628$, $p = .059$, partial eta-squared = .123, were not found to be significant. Post-hoc comparisons revealed that “Forward” accuracy did not differ from the “Reverse” condition ($t(30) = -.045$, $p = 0.91$, Cohen’s $d = -.02$), but that “Forward” accuracy was significantly lower than performance in both slideshow conditions (Slideshow: $t(30) = -2.65$, $p = 0.013$, Cohen’s $d = -.94$); Blank-

Slideshow: $t(30)=-2.56$, $p=0.016$, Cohen's $d=-.91$) as shown in Figure 3. Finally, one-sample t-tests revealed that performance in all 4 conditions was above chance ($p<0.001$ for all tests).

Discussion

Participants in Experiment 1 could accurately categorize dyadic familiarity in all stimulus conditions, but performed significantly better in both static-image slideshow conditions, suggesting that in this case, the dynamic stimulus is not superior for thin-slice recognition. While motion is known to enhance performance across certain thin-slice tasks, our results suggest that motion may diminish the ability to categorize familiarity. While dynamic information improves thin-slice perception in some cases, judgments utilizing only static information can be successfully executed. Male sexual orientation can be determined from a static image presented for only 50 ms (Rule & Ambady, 2008), and the sexual orientation of female targets can be perceived from a series of eight still photographs (Ambady et al., 1999). Specific trait inferences, such as trustworthiness, competence, and likeability, can be determined from unfamiliar still photographs in as little as 100 ms (Willis & Todorov, 2006). However, while previous results have demonstrated that thin-slice perception *can* be accomplished with static images, to our knowledge this is the first report that static images may actually be better in some circumstances than a full dynamic sequence. We continued in Experiment 2 to further examine the contribution of static frames and motion to thin-slice perception by manipulating the speed of video playback.

Experiment 2

To further explore how motion properties affect thin-slice perception in this task, we examined how video speed influenced familiarity recognition. If natural motion is especially useful for thin-slice judgments, we predicted that both faster and slower video playback would likely impair performance, since in both cases natural gesture is disrupted. Alternatively, if static frames are particularly important for recognition of mutual familiarity, faster playback should be especially bad (since less time per frame is available) and slower playback should improve performance (since more time per frame is available).

Method

Thirty-six young adults (7F/11M in the fast video condition, 10F/8M in the slow video condition) in the age range of 18-34 years participated in Experiment 2. The stimuli were the same as those used in Experiment 1, except we manipulated the video playback rate. In the slow video condition, participants viewed the videos at original speed and at a reduced rate (0.5×). In the fast video condition, videos were viewed at an increased rate (2×) and at the original speed, which allowed us to look at cost directly. The order in which original and speed-manipulated videos were viewed was counterbalanced.

Results

We analyzed accuracy by calculating the difference scores (number correct for regular speed videos minus number correct for speed manipulated videos) for both the fast and slow

conditions to evaluate whether there was a cost associated when viewing speed was manipulated compared to regular speed videos. We conducted a $2 \times 2 \times 2$ between-subjects ANOVA with sex (female or male), cost type (fast or slow), and task order as the between subjects factors. We included task order in our analysis to account for variability associated with a practice effect resulting from exposure to the same videos in both the regular speed and speed manipulated conditions. This analysis revealed a main effect of video speed, ($F(1,28) = 4.37, p = .046$, partial eta-squared = .135). The mean difference score for the fast condition ($M = 1.90, SE = 1.06$) was positive, which indicated participants performed better when the videos were played at regular speed, whereas the mean difference score for the slow condition ($M = -1.38, SE = 1.15$) was negative and indicated participants were more accurate in the slow versus regular condition. Participant sex ($F(1,28) = .064, p = .802$, partial eta-squared = .002) and task order ($F(1,28) = .004, p = .947$, partial eta-squared = .001) were not found to be significant, nor were there any significant interactions. Paired samples t-tests demonstrated that within-subject performance in the slow and regular conditions did not significantly differ from one another $t(17) = .559, p = .584$, Cohen's $d = .13$, nor did within-subject performance on fast and regular videos differ, $t(17) = -1.627, p = .122$, Cohen's $d = -.39$. That is, neither manipulation significantly affected performance considered singly, but the cost induced by each was statistically different.

We also observed a marginally significant between-group difference in accuracy for regular-speed videos $t(34) = -1.972, p = .057$, Cohen's $d = -.66$ in our two participant groups. Finally, to assess the extent to which slow videos behaved like slideshows, we compared performance in the slow condition of this task to performance in the "Slideshow" condition of Experiment 1. Participants performed slightly better in the static-slideshow condition from Exp. 1 ($M = 21.44, SE = .77$) than the slow condition ($M = 19.83, SE = .85$), although this was not found to be significant $t(32) = 1.39, p = .174$, Cohen's $d = .48$. See Figure 4 for accuracy across all conditions.

General Discussion

In Experiment 1, optimal performance for the detection of personal familiarity was obtained with the use of static frames with and without a blank screen in between images whereas in Experiment 2, there was evidence of a larger performance cost when video speed was sped up then when it was slowed down. Together our results suggest that, contrary to previous thin-slice results, observers' ability to accurately extract information from static frames may be critically important for detecting mutual familiarity. In Experiment 1, participants performed similarly when the video was played in the forward direction versus the reverse direction, indicating that the disruption of temporal order did not significantly impair familiarity judgments. By itself, this suggests that the preservation of the static frames in a sequence is sufficient for performance to be maintained in this task. Temporal order does influence the ability to determine whether an individual is alone or engaged in an interaction (Balas et al., 2012), suggesting that in other thin-slice tasks, the disruption of natural movement by video reversal limits the efficient encoding of diagnostic information. Here, natural forward movement is either not used at all, or used in a fairly coarse manner that does not take into account the order of events or the direction of gestures. The observation that static image slideshows (with or without apparent motion between frames) lead to

increased performance in Experiment 1 further suggests that individual frames may be an important locus of diagnostic information in our task, which is supported by the more subtle observed effects of video speed in Experiment 2. These results thus suggest that at least in some cases, natural motion may not be as useful as static appearance for thin-slice perception, which poses a bit of a puzzle: Why should static frames be especially useful for some judgments?

Thin-slice perception may rely on a number of component processes, including emotion recognition, face recognition, and object recognition – all of which are known to benefit from dynamic information. The ability to detect the emotional state of a person engaged in an interaction relies heavily on specific facial expressions, and those expressions are thought to be intrinsically dynamic in nature (de la Rosa et al., 2013). Identity, gender and emotional state can be perceived from biological motion of the body (Johansson, 1973) or the face (Bassili, 1978). More general affective states have also been assessed within the thin-slice domain (Hall & Bernieri, 2001). Positive and negative affect can be detected from thin-slices of varying lengths from dyadic interactions (Carney, Colvin, & Hall, 2007) and motion appears to play a role in these tasks as well. Considering face recognition more broadly, motion appears to confer a benefit on person recognition (O'Toole, Roark, & Abdi, 2002). The behavioral recognition advantage for dynamic faces is especially strong when face images are degraded (Lander & Bruce, 2000), but obtains more generally as well (Thornton & Kourtzi, 2002). This overview of how motion impacts these presumably related visual processes is far from exhaustive, but nonetheless demonstrates that dynamic information generally contributes to enhanced recognition performance across a variety of stimuli and tasks.

How then to explain the advantage that observers exhibited for observing our static slideshow stimuli in Experiment 1? If motion benefits other thin-slice judgments as well as other recognition tasks that we presume are related to thin-slice perception, how can we interpret the results we obtained in our study? We offer one account of this counter-intuitive result that we suggest follows naturally from some of the known properties of face and emotion processing. It may be the case that natural motion inhibits access to transient static features of the scene and limits the ability to distinguish specific static expressions or gestures needed to determine the personal familiarity of a dyad. For example, the frozen face effect (FFE) refers to the finding that static images taken from a video of someone speaking usually are rated as less flattering than the video itself, which is thought to be due to the lack of motion (Post, Haberman, Iwaki, & Whitney, 2012). Given that the continuous motion of faces in this context serves to hinder the encoding of static images within the sequence, we suggest that a similar phenomenon may contribute to the slideshow advantage we observed. In particular, “microexpressions” – fleeting facial gestures – are known to be both diagnostic of complex mental states including deception and are largely involuntary and difficult to suppress (Porter, ten Brinke, & Wallace, 2012). If “leakage” of microexpressions during interaction with familiar and unfamiliar individuals provides diagnostic information about that interaction, then presenting them outside of a motion sequence may be useful to the observer. Training observers with microexpression recognition impacts their subsequent performance in real-world social perception tasks (Matsumoto & Hwang, 2011), suggesting that there is indeed an important link between the ability to measure transient appearance

and the ability to perform thin-slice tasks effectively. While we cannot unequivocally state that specific microexpressions provide the basis for our effects, this account of our data makes several interesting predictions. First, if it is true that there are specific microexpressions that provide highly diagnostic information, it should be possible to identify them. An extended analysis of our own stimuli (and others) with an emphasis on determining which static frames are critical for recognition may help support our proposal that specific instants within a larger sequence carry most of the information for recognition. Another possibility that we have not addressed in the current study is whether it is the reduction of our sequences to a sparse set of images that benefits recognition, or whether it is the reduction of those sequences to a sparse set of *transitions*. While reduced apparent motion did not decrease accuracy in the “Blank Slideshow” condition, the order of static frames was preserved when we sampled them from the original sequences in Experiment 1, which may mean that some comparison between sequentially presented images contributes to performance.

Finally, it would be interesting to revisit a number of ecologically relevant thin-slice tasks to compare static and dynamic conditions directly. The relative contribution of motion vs. frames may be an important way to determine the number of distinct mechanisms that contribute to these highly complex social inferences. Presently, the current results suggest that the ability to judge personal familiarity is enhanced when motion is substantially decreased, which to our knowledge has not been observed in the context of thin-slice perception. Future work to disentangle and identify the critical cues necessary for executing judgments of interactions from thin slices of behavior will certainly yield important insights into the computational and behavioral basis of these challenging recognition tasks.

Acknowledgments

BB was supported by NIGMS Grant #P20 103505. Thanks also to Dan Gu for technical support.

References

- Abramovitch R. Children’s recognition of situational aspects of facial expression. *Child Development*. 1977; 48:459–463.
- Ambadar Z, Schooler JW, Cohn JF. Deciphering the enigmatic face the importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*. 2005; 16(5):403–410.
- Ambady N, Conner B, Hallahan E. Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*. 1999; 77(3):538–547. [PubMed: 10510507]
- Ambady N, Gray HM. On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology*. 2002; 83(4):947–961. [PubMed: 12374446]
- Ambady N, Krabbenhoft M, Hogan D. The 30-sec sale: Using thin slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology*. 2006; 16(1):4–13.
- Ambady N, Rosenthal R. Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*. 1993; 64(3): 431–441.
- Balas B, Kanwisher N, Saxe R. Thin-slice perception develops slowly. *Journal of Experimental Child Psychology*. 2012; 112:257–264. [PubMed: 22417920]

- Bassili JN. Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*. 1978; 4(3):373–379. [PubMed: 681886]
- Benjamin GR, Creider CA. Social distinctions in nonverbal behavior. *Semiotica*. 1975; 14:52–60.
- Brainard DH. The psychophysics toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
- Carney DR, Colvin CR, Hall JA. A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*. 2007; 41(5):1054–1072.
- Costanzo M, Archer D. Interpreting the expressive behavior of others: The interpersonal perception task. *Journal of Nonverbal Behavior*. 1989; 13(4):225–245.
- de la Rosa S, Giese M, Bülthoff HH, Curio C. The contribution of different cues of facial movement to the emotional facial expression adaptation aftereffect. *Journal of vision*. 2013; 13(1):1–15.
- DePaulo BM. Nonverbal behavior and self-presentation. *Psychological Bulletin*. 1992; 111(2):203–243. [PubMed: 1557474]
- Feyereisen P. The behavioral cues of familiarity during social interactions among human adults: A review of the literature and some observations in normal and demented elderly subjects. *Behavioural Processes*. 1994; 33:189–212. [PubMed: 24925246]
- Hall, JA.; Bernieri, FJ. *Interpersonal sensitivity: Theory and measurement*. Erlbaum; NJ: 2001.
- Ickes, W. Measuring empathic accuracy. In: Hall, JA.; Bernieri, FJ., editors. *Interpersonal sensitivity: Theory and measurement*. Erlbaum; Mahwah, N.J.: 2001. p. 219–241.
- Johansson G. Visual perception of biological motion and a model for its analysis. *Attention, Perception, & Psychophysics*. 1973; 14(2):201–211.
- Lander K, Bruce V. Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*. 2000; 12:259–272.
- Matsumoto D, Hwang HS. Evidence for training the ability to read microexpressions of emotion. *Motivation and Emotion*. 2011; 35:181–191.
- O'Toole AJ, Roark D, Abdi H. Recognition of moving faces: A psychological and neural framework. *Trends in Cognitive Science*. 2002; 6:261–266.
- Porter S, ten Brinke L, Wallace B. Secrets and Lies; Involuntary Leakage in Deceptive Facial Expressions as a Function of Emotional Intensity. *Journal of Nonverbal Behavior*. 2012; 36:23–37.
- Post RB, Haberman J, Iwaki L, Whitney D. The frozen face effect: why static photographs may not do you justice. *Frontiers in Psychology*. 2012; 3(22):1–7. [PubMed: 22279440]
- Richeson JA, Shelton NJ. Brief report: Thin slices of racial bias. *Journal of Nonverbal Behavior*. 2005; 29(1):75–86.
- Rule NO, Ambady N. Brief exposures: Male sexual orientation is accurately perceived at 50 ms. *Journal of Experimental Social Psychology*. 2008; 44(4):1100–1105.
- Schiff W, Banka L, de Bordes Galdi G. Recognizing people seen in events via dynamic “mug shots”. *The American Journal of Psychology*. 1986; 99(2):219–231. [PubMed: 3766815]
- Thornton IM, Kourtzi Z. A matching advantage for dynamic human faces. *Perception*. 2002; 31(1): 113–132. [PubMed: 11922118]
- Willis J, Todorov A. First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*. 2006; 17(7):592–598. [PubMed: 16866745]

Figure 1.
Example of trial containing familiar and unfamiliar interactions.

Figure 2.
Example of an abbreviated static-image slideshow sequence.

Figure 3.
Proportion correct for each condition of Experiment 1.

Figure 4.
Proportion correct for each condition of Experiment 2.