YouTube as a datasource for nonverbal communication research

Baxter, P.A.; Vromans, Ruben; Swerts, Marc

YouTube as a Datasource for Nonverbal Communication Research

Peta A. Baxter, Ruben D. Vromans, and Marc G. J. Swerts

Tilburg University

Peta A. Baxter, Department of Communication and Information Sciences, Tilburg University (p.a.baxter@tilburguniversity.edu); Ruben D. Vromans, Department of Communication and Information Sciences, Tilburg University (r.d.vromans@tilburguniversity.edu); Marc G. J. Swerts, Tilburg center for Cognition and Communication, Tilburg University (m.g.j.swerts@tilburguniversity.edu).

Correspondence concerning this extended abstract should be addressed to Peta Baxter, Department of Communication and Information Sciences, Tilburg University, 5000 LE, Tilburg, The Netherlands.

Word count: 750 (excluding references)

Introduction

A key function of nonverbal communication is expressing emotions (Argyle, 2013). However, obtaining high ecological validity remains a challenge in studying the perception of nonverbal expressions of emotions. Indeed, emotion research typically uses stimuli depicting a stereotypic nonverbal expression of an emotion, thereby disregarding the fact that in real life emotions are gradient and can be felt and expressed with different degrees (Hess, Blairy, & Kleck, 1997). In addition, although individuals normally use several body parts (often simultaneously) to express emotions, e.g. using both facial expressions and gestures, whole-body combinations are rarely considered in research (cf. Bänziger & Scherer, 2007; Scherer & Ellgring, 2007; Van den Stock et al., 2007). To best represent the spontaneous fluctuations and whole-body nature of expressed emotions, stimuli consisting of naturalistic spontaneous portrayals should be prioritized over acted ones (Fernández-Dols & Crivelli, 2013; Scherer & Ellgring, 2007; Wilting, Krahmer, & Swerts, 2006).

However, developing naturalistic stimuli is challenging because spontaneous emotional reactions are difficult to elicit and capture in a controlled manner (e.g., in comparable situational contexts). A promising solution is the platform YouTube, which contains a large amount of video data of people expressing emotions in similar contexts. It is open access and offers insight into the (cross-cultural) behaviors of individuals. Moreover, semi-controlled settings can be artificially created by using data in fixed settings (e.g., game shows, political debates). Some research has already used YouTube data to analyze the nonverbal behavior of vloggers (Biel & Gatica-Perez, 2010). However, no study has explicitly sought to determine whether YouTube could be a suitable source of whole-body, naturalistic data for the study of nonverbal expressions of emotion. We hypothesized that (1) YouTube

contains data in which observers can perceive different levels of expressed affective intensity (2) and that observers use multiple body parts to do so.

Method

Stimuli

To develop naturalistic stimuli depicting different levels of affective intensity, we used clips of the TV game show "Who wants to be a millionaire?". Specifically, clips were selected to show the moment immediately after nine different contestants had secured 1.000$ (low affective intensity), 32.000$ (medium affective intensity) or 1.000.000$ (high affective intensity). The clips were comparable in terms of duration (3s on average), focused on the contestants' upper body, and were muted so that auditory cues would not be accessible to the observers (see Figure 1). This resulted in a total of 27 clips: 9 (contestants) × 3 (levels of affective intensity). This created a semi-controlled setting: indeed, all contestants were undergoing comparable affective experiences that were induced by the fixed setup of the show. In this case, the different stages of monetary reward served as an ideal operationalization of the levels of affective intensity.

Perception test

In total, 112 participants (74 females, Mage = 28.50 [SD = 10.50], age range = 19-68) took part in an online perception test. Each trial consisted of a fixation cross (2s), followed by the clip (3s). After each clip, a first question addressed the perceived level of expressiveness and was judged on a 4-point scale. Then, participants were asked to indicate which body parts were most informative for assessing levels of expressiveness. The answer options comprised the face/head, hand/arms, torso, whereby participants could select all that applies.
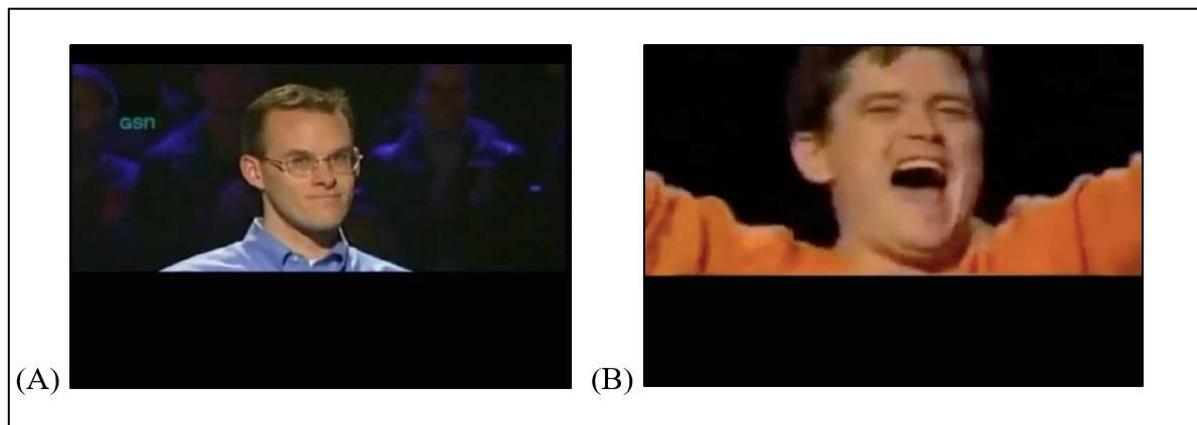
Figure 1. Still images of the stimuli used in the perception study displaying low-level of affective intensity (A) and high-level of affective intensity (B).

Results

Perceived level of expressiveness

A one-way repeated measure ANOVA revealed a significant main effect of level of affective intensity on the perceived level of expressiveness, $F(2, 222) = 312.30$, $p < .001$, $\eta^2 = .74$. Repeated contrasts showed that the level of perceived expressiveness in high intensive situations ($M = 3.20$, $SD = 0.49$, 95% CI [3.21, 3.33]) was higher than on a medium intensive situation ($M = 2.52$, $SD = 0.56$, 95% CI [2.41, 2.62]), $p < .001$. The level of perceived expressiveness in the medium intensive situations, in turn, was also higher than in the low intensive situations ($M = 1.69$, $SD = .63$, 95% CI [1.58, 1.81]), $p < .001$.

Body parts

Table 1 represents a matrix that depicts the different body parts indicated as being most expressive by the respondents in the different levels of affective intensity. Visual inspection of the table suggests that the face/head is the most expressive physical attribute as compared to the other ones. Moreover, as affective intensity increases, the percentages of all

other body parts indicated as expressive increase as well, demonstrating the importance of whole-body cues in assessing affective intensity.

Table 1

Body Parts Indicated as Most Expressive (in percentages)

Level of affective intensity	Face/Head	Arms/Hands	Torso
Low	67.99	0.32	3.73
Medium	84.59	19.17	10.31
High	84.91	41.46	17.40

Conclusion

In this study we show that YouTube can be a suitable source of data for investigating spontaneous and whole-body nonverbal expressions of emotion. From the stimuli, observers were able to perceive different levels of expressed affective intensity by using whole-body cues. In conclusion, the use of YouTube as datasource can be considered to improve the ecological validity of nonverbal communication research.

References

Argyle, M. (2013). *Bodily communication*. Routledge.

Bänziger, T., & Scherer, K. R. (2007, September). Using actor portrayals to systematically study multimodal emotion expression: The GEMEP corpus. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 476-487). Springer Berlin Heidelberg.

Biel, J. I., & Gatica-Perez, D. (2011). VlogSense: Conversational behavior and social attention in YouTube. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 7, 33.

Fernández-Dols, J. M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, 5, 24-29.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*, 241-257.

Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns?. *Emotion*, *7*, 158.

Van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*, 487.

Wilting, J., Krahmer, E. J., & Swerts, M. G. J. (2006, September). Real vs. acted emotional speech. In *Proceedings of the international conference on spoken language processing (interspeech 2006)*, Pittsburgh, PA, USA.