THE ROLE OF PRESENTER ACTIONS IN THE EVALUATION OF ORAL PRESENTATIONS: EVIDENCE FROM FMRI AND MOTION-VECTOR ANALYSIS

Ryuta ISEKI¹⁾²⁾, Saori KAJIMA¹⁾, Xiangyue ZHAO¹⁾, Xuefeng LIANG¹⁾, and Satoshi TSUJIMOTO¹⁾³⁾

¹⁾*Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Japan*

²⁾*Department of Human Sciences, Taisho University, Japan*

³⁾*The Nielsen Company Singapore Pte Ltd, Singapore*

The present study aimed to explore the neurocognitive mechanisms underlying the evaluation of oral presentations. The top 10 and bottom 10 TED presentations were selected from over 2,000 TED talks based on viewer ratings. Thin slices from these 20 videos were used as stimuli. The first experiment showed that the original ratings could be represented by evaluations of 30-sec clips. In the second experiment, fMRI BOLD signals from viewers were examined while they watched the thin-sliced TED talks. The results showed significantly greater activation in the parts of the occipitotemporal cortex related to visual motion and social cognition while watching the top 10 clips than when watching the bottom 10 clips. In the third experiment, the magnitude of motion in the video clips, as quantified by motion vectors, differed significantly between two distinct presentation groups. These results provide comprehensive evidence for the importance of presenter actions for oral presentations to be evaluated highly by audiences.

Key words: thin slicing, extrastriate body area, EBA, video analysis, free viewing

INTRODUCTION

Effective oral presentation is often required for success in various professional fields and academic settings (Murphy, Hildebrandt, & Thomas, 1997). Therefore, there is a growing need for effective educational methods to teach these presentation skills. However, the factors contributing to the evaluation of presentations and their underlying neurocognitive mechanisms are not yet fully understood.

A key approach to this issue may be the thin-slicing technique. Thin slicing refers to the judgment based on an exposure to a narrow time-window of experience. Since the original report by Ambady and Rosenthal (1992), many studies have shown that consensual judgments of others are unexpectedly accurate even when they are based on a brief exposure

This study was conducted using the MRI scanner and related facilities of Kokoro Research Center, Kyoto University. We thank Ms. Maki Terao for her technical support on MRI data acquisition.

This work was financially supported by the ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan) and Grants-in-Aid from JSPS (15K12049 and 26282218).

Correspondence concerning this article should be addressed to Satoshi Tsujimoto, Nielsen Consumer Neuroscience, The Nielsen Company Singapore Pte LTD, Goldbell Towers, 47 Scotts Road #13-00, Singapore 228233 (e-mail: tsujimoto.gml@gmail.com).

to the targets' behavior. For example, Ambady and Rosenthal (1993) reported that the viewers' ratings of teachers based on brief, silent clips less than 30 sec long were significantly correlated with the end-of-semester student evaluation for those teachers. Likewise, judgements based on thin-sliced stimuli have significantly predicted the performance of salespeople (Ambady, Krabbenhoft, & Hogan, 2006), intellectual abilities (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004), personality (Borkenau, Brecke, Möttig, & Paelecke, 2009), and the level of language development (Walton & Ingersoll, 2016). Thus, in the context of the evaluation of oral presentations, nonverbal aspects are likely to be critical for the presentations to be highly evaluated; therefore, the thin-slicing technique will be useful for the exploration of these aspects.

The purpose of this study is to reveal the mechanisms behind the evaluation of oral presentation based on its nonverbal aspects. We used thin slices of TED (Technology, Entertainment, Design; see <https://www.ted.com/talks>) talks as stimuli. Over 2,000 TED presentations with considerably similar formats are available online and viewers can send feedback by rating each presentation online after watching them. We conducted three experiments using these stimuli, each of which allowed us to approach our research purpose from different aspects, i.e., behavior, brain, and video analyses, for Experiments 1, 2, and 3, respectively.

In Experiment 1, we selected the top 10 and bottom 10 TED videos based on their ratings and tested whether thin slices of those videos could be evaluated similarly to their original ratings. To explore the neural mechanisms related to the evaluation of thin-sliced oral presentations, Experiment 2 compared the viewers' blood oxygenation level-dependent (BOLD) brain activity using functional magnetic resonance imaging (fMRI) while they watched those thin-sliced clips. Finally, in Experiment 3, the motion vectors for each pixel were used to analyze the magnitude of motion across the video frames in those clips. By integrating these three experiments, we provide comprehensive evidence on the fundamental role of presenter actions in the high evaluation of oral presentations by audiences.

EXPERIMENT 1

The aim of this experiment is to test the feasibility of thin slicing in oral presentations. The experiment also aimed at exploring the reason behind the evaluation of thin-sliced presentation clips at the level of questionnaires.

METHOD

Participants

Forty-three paid volunteers were recruited as participants. All participants had normal or corrected-to-normal visual acuity. Written informed consent was obtained prior to the commencement of the experiment. Participants were randomly assigned into two different groups: Group 1 consisted of 11 women and 11 men aged between 19 and 33 years, and Group 2 were 11 women and 10 men aged between 19 to 29 years. The Ethics Committee of Kokoro Research Center, Kyoto University, approved the study.

Fig. 1. Distribution of 2,027 TED talks against the rating index. The index was calculated based on viewers’ ratings (see the Methods for details) and was transformed into a log scale. Darker gray bars indicate the groups that included the selected videos.

Stimuli

The stimuli were 30-sec video clips prepared from TED conference presentations. The TED video archive is available online under the Creative Commons licenses. After viewing these online videos, viewers are allowed to provide their ratings according to 14 attributes, such as informative, fascinating, and inspiring. At the time of this experiment (August 14, 2015), 2,027 presentations were listed in the online archive and 4,575,411 ratings had been provided.

To select presentations with high and low evaluations, we focused on the following four of the 14 attributes, inspiring, informative, persuasive, and fascinating, which had received a relatively large number of viewer evaluations. The video-to-video ratings among these four attributes were also highly correlated to each other with a correlation coefficient of 0.58–0.78. Thus, these attributes were suitable for a single representative index of viewers’ ratings. The index was calculated using the following formula:

$$\text{Index} = (\text{Inspiring} + \text{Informative} + \text{Persuasive} + \text{Fascinating})/4$$

As shown in Fig. 1, the index values for the 2,027 presentations were distributed approximately log-normally. We selected the top 10 and bottom 10 TED videos as high-evaluation and low-evaluation presentations (Hp and Lp, respectively) based on this rating (Fig. 1).

Each 30-sec duration thin-sliced video clip was selected from these 20 clips. A volunteer who was naïve to the experiment extracted scenes according to the following criteria: a scene when she felt climax, and a scene exclusively showing the presenter (image of audience not included) without audiences’ laughter and applause.

The audio was not manipulated. The language used in all presentations was English, with no subtitles in either Japanese or English.

Procedure

The experiment was controlled with PsychoPy (1.83) software running on a Windows PC. The video clips were shown on a 17.3-inch screen and participants listened to audio through a headphone.

Prior to the experimental sessions, each participant was instructed to give his/her ratings following the viewing of the presentation video. Ratings were performed using 5-point Likert scales (higher values mean a more-positive evaluation) and responses were made using a numeric keypad. Participants were also told that the presentation is given in English and that they do not have to understand the contents.

A practice session followed the instruction. The experimental sessions then commenced. Twenty 30-sec video clips were presented to participants in random order and each participant gave their rating after viewing each video.

The rated items differed between participant groups 1 and 2. Group 1 rated two items: (1) overall rating of the video (“how good do you think the presentation was?”) and (2) the confidence of the rating (“how confident were you on this rating?”). Group 2 rated five presentation attributes: activeness, attractiveness, passion, persuasiveness, and estimated rank in the original TED archive ranking.

Data analysis

Ratings for each item were separately averaged for the 10 Hp and 10 Lp videos observed by each participant. The grand-averaged scores were then calculated between participants for each item separately. Paired t-test was used to statistically compare the scores for each item between Hp and Lp videos. We also computed Cohen’s *d* to report effect size.

RESULTS AND DISCUSSION

Group 1 gave significantly higher overall ratings to Hp than to Lp videos (Hp, 3.67; Lp, 3.06; $t = 6.84$, $p < .001$, Cohen’s $d = 1.49$). The confidence of this rating was only around the median (3 out of 5) and was not significantly different between Hp and Lp videos (Hp, 3.26; Lp, 3.65; $t = 1.64$, $p = .12$, Cohen’s $d = .36$).

A full TED talk typically lasts 20 min and the current video stimuli were 30-sec thin slices of the talks. The distinction between Hp and Lp based on ratings of full-length talks (an over 4.5 million dataset) was nevertheless represented by a thin-slice-based evaluation of just 22. Together with the difficulty of comprehending a second language without subtitles, it is unlikely that ratings were based on the contents of the talk. Thus, the results indicate that the thin-slicing method and theory (Ambady et al., 2006; Ambady & Rosenthal, 1993; Borkenau et al., 2004) can be applied to the field of oral presentations.

To explore the possible reasons for the distinct overall ratings, we then analyzed the Group 2 ratings. The results are summarized in Table 1. As shown, the ratings were consistently higher in Hp than Lp for all tested attributes. The result shows that the thin-slicing effect is robust. Notably, Hp videos were evaluated as being more “persuasive” than Lp videos. Considering the nature of the current stimuli, it is unlikely that the viewers were actually persuaded by the contents of the presentation. This is consistent with previous findings that showed that the style of the lecturer, such as gestures and voice tone, instead of the content itself has a considerable impact on university class evaluations, including those regarding textbook quality and fairness of academic assessment (Williams & Ceci, 1997).

Table 1. Questionnaire results for Group 2 participants in Experiment 1

	Hp	Lp	<i>t</i>	<i>p</i>	Cohen’s <i>d</i>
Activeness	3.56	2.74	12.9	<.001	2.82
Attractiveness	3.41	2.90	6.77	<.001	1.48
Passion	3.52	3.00	5.46	<.001	1.19
Persuasiveness	3.58	3.16	7.80	<.001	1.70
Rank	3.66	3.09	8.53	<.001	1.86

In summary, these results indicate that the thin-slicing technique can be applied effectively to the oral presentation settings. In addition, the Group 2 results show that articulated evaluation via questionnaires is not necessarily useful for assessing the reasons and mechanisms behind the evaluation of oral presentation. Therefore, we performed Experiment 2 in an attempt to explore those issues by focusing on the brain response to the thin-sliced oral presentations.

EXPERIMENT 2

In Experiment 1, we showed that the thin-slicing technique is applicable to the evaluation of TED talks. However, the questionnaire-based assessment was also shown to be insufficient to reveal the reason behind the evaluation. Experiment 2 was designed to explore this issue.

METHOD

Participants

Thirty young adults participated in this experiment. None of them had participated in the first experiment. They were divided into two groups: Group A and Group B. Group A consisted of 20 right-handed normal adults (9 women, aged 20–43 years), and the remaining 10 right-handed normal adults were classified into Group B (4 women, aged 20–39 years). All participants had no history of neurological diseases and had normal or corrected-to-normal visual acuity. Written informed consent was obtained from all participants. The Ethics Committee of Kokoro Research Center, Kyoto University, approved the study.

Stimuli and participant tasks

The 20 thin-slice clips used in Experiment 1, i.e., the 10 Hp and 10 Lp video clips, were also used in this experiment. Of these video clips, the videos ranked by odd numbers (e.g., top 1, 3, 5, 7, 9 and bottom 1, 3, 5, 7, 9) were presented to Group A participants, and the even-ranked videos were shown to Group B.

In addition, a set of 10 clips were newly prepared for this experiment from the middle-ranked videos in the distribution in Fig. 1. This set of clips was used as the control presentation (Cp). The length of Cp video clips was also 30 sec, but the style differed from that of Hp and Lp. Three static images were randomly extracted from each presentation and those images were sequentially presented for 10 sec each. The initial and final second of each 10-sec image was a fade-in or -out to smooth the transitions between images while related audio was played in reverse. Thus, no meaning was conveyed by the audio, with auditory stimulation being kept comparable.

In addition to the “Rest” period of 30 sec, a sequential set (blocks) was composed in the order of either Hp, Cp, Lp, Rest or Lp, Cp, Hp, Rest. Five such blocks were presented to each participant while their brain was being scanned (Fig. 2). These two types of blocks were alternately interchanged using the five sets. The presentation orders of the five sets were randomized between participants. During Hp, Cp, and Lp, the video clip for each condition was presented and only the fixation cross was presented at the center of the black screen during the Rest period. Each block lasted 2 min (30 sec $\times$ 4 conditions; see Fig. 2); therefore, the total scan length was 10 min.

The participant’s task was to press the key every time when the presenter in the video changed from one to the other. The instruction was given to participants using a demo set of 2 min and they were informed that they do not need to understand the contents.

The visual stimuli were presented on an MRI-compatible display (Nordic Neuro Lab, Inc., Bergen, Norway) and participants viewed the stimuli through a mirror attached to the head coil. The audio was delivered via MRI-compatible headphones. The responses were recorded by a fiber-optic response device

Fig. 2. Sequence of a representative block for Experiment 2. During Hp, Lp, and Cp conditions, the thin-sliced video clips were presented to the participant. During the rest period, a white fixation cross appeared on the black background of the screen. Each participant was exposed to five such blocks during an fMRI scan. The order of Hp and Lp conditions was alternately exchanged across blocks, with Cp and rest conditions being fixed. **Abbreviations:** Hp, high-evaluated presentation; Lp, low-evaluated presentation; Cp, control presentation.

(Current Designs, Inc., Philadelphia, PA, USA). The presentation of audiovisual stimuli and collection of key-press responses were controlled by Octave 4.0.0 with Psychtoolbox 3.0.12 (Brainard, 1997; Pelli, 1997) on a Mackintosh computer.

MRI data acquisition

All the MRI data were collected using a Siemens 3.0T MAGNETOM Verio scanner (Siemens, Germany) with a 32-channel head coil in the Kokoro Research Center, Kyoto University. Following the T_1 -weighted sagittal localizer scanning, the functional images were acquired using a gradient-echo echo-planar imaging (EPI) sequence, which are sensitive to BOLD contrast. The sequence parameters were as follows: TR = 3 s, TE = 25 ms, flip angle = 80° , FOV = $192 \text{ mm} \times 192 \text{ mm}$, slice thickness = 3 mm, and 47 horizontal slices. Then, high-resolution structural images were collected using a T_1 -weighted magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) pulse sequence (TR = 1900 mm, TE = 2.52 ms, FOV = $256 \text{ mm} \times 256 \text{ mm}$, slice thickness = 1 mm, 192 slices).

Data analysis

The preprocessing and statistical analyses were performed using Statistical Parametric Mapping 12 (SPM 12, Wellcome Trust Centre for Neuroimaging, London, UK), which was implemented in MATLAB 9.0.

Functional images for the initial four scans were discarded for the preprocessing. The remaining images were then corrected for head motion and for slice-acquisition time differences. The resulting data sets were coregistered to the T_1 -weighted MPRAGE and spatially normalized into the Montreal Neurological Institute (MNI) template, followed by a spatial smoothing using a Gaussian Kernel with a Full-width at half-maximum (FWHM) of 8 mm.

The preprocessed images were then statistically analyzed in two steps at individual and group levels. At the individual level, the data were modeled using the general linear model principles. The design matrix comprised two clip conditions (i.e., Hp vs. Lp, Hp vs. Cp, and Lp vs. Cp) modeled using a square wave function convolved with the canonical hemodynamic response function. The movement parameters taken during the motion correction stage were also included as covariates in the model. At the group level, the condition-related contrast images captured in the preceding individual-level analysis were analyzed using a random-effect model.

RESULTS AND DISCUSSION

First, to examine brain responses to the oral presentation viewing, we compared the BOLD signals of Group A participants during Hp and Lp conditions with those during the Cp condition. Fig. 3A and 3B separately show the results of each comparison on an inflated brain. The activation patterns were quite similar in both comparisons. According to our criteria ($p < .001$, uncorrected), the bilateral occipitotemporal cortex showed significantly higher activation in Hp and Lp than in Cp conditions. In the left hemisphere, the activated areas were extended to anterior parts of the superior temporal gyrus (STG). In addition, activation was also observed in the left frontal eye field.

The fact that we observed a similar pattern of activation for both Hp and Lp conditions is reasonable because both Hp and Lp are derived from the TED talks, which are based on a common format. The activation in the occipitotemporal cortex as well as the left STG has been reported repeatedly in previous studies using dynamic audiovisual stimuli for a longer duration than thin slicing (Bordier, Puja, & Macaluso, 2013). At the behavioral level, thin-slice experiments have been shown to dominate the visual domain (Elisha Babad, Avni-Babad, & Rosenthal, 2003; Borkenau et al., 2004). The present results support these findings at the level of the BOLD response.

The occipitotemporal clusters included the V5/MT area, which is primarily related to motion-related vision. Thus, this area would reflect the motion in the videos. The frontal eye fields (FEF) activation also supports this interpretation.

To examine whether the level of the activation differs between Hp and Lp conditions, we compared their BOLD signals. We found three clusters with significant effects ($p < .001$, uncorrected) in the left hemisphere (Fig. 4), V2/V3d, middle occipital gyrus (MOG; BA37), and STG. Contrarily, we did not observe any clusters in the contrast of Lp over Hp.

Anatomically, the V2/V3d cluster is thought to be a part of the dorsal stream (Kujovic et al., 2013). This area receives inputs from V1 via V2, and sends the projection to the parietal cortex. Although the functional role of this cluster is still unclear, it appears to be involved in the processing of global motion (Braddick et al., 2001). The MOG cluster included the extrastriate body area (EBA), which is thought to play a role in perception of people's bodies (Downing, Jiang, Shuman, & Kanwisher, 2001). EBA activation has been reported in relation to the observation of other people's purposeful actions such as pointing, grasping, and basketball-shooting actions (J. W. Kable & Chatterjee, 2006; Kable, Kan, Wilson, Thompson-Schill, & Chatterjee, 2005; Pierno et al., 2009). The third cluster, STG, has been reported to be involved in social cognition (Bigler et al., 2007; Pehrs et al., 2017). Pehrs et al. (2017) showed that the top-down signal from the temporal pole modulates this area.

To test whether these findings can be generalized to other sets of TED talks, we analyzed the BOLD signals from Group B participants, who viewed different sets of Hp and Lp videos to the other group of participants. In this analysis, we particularly focused on EBA and used region of interest (ROI) analysis. The ROI was set over the EBA, which was over the spherical regions centered on the MNI coordinates [51, -71, 1] based on the

Fig. 3. Blood oxygenation level-dependent (BOLD) responses related to the viewing of oral presentation. **A:** Activation maps on an inflated brain showing voxels with significantly greater BOLD responses to high-evaluated presentations than to control presentations ($p < .001$, uncorrected). **B:** The same plot as in panel A, but for the BOLD responses to low-evaluated presentations compared with control presentations.

Fig. 4. Activation maps on an inflated brain showing clusters with greater BOLD responses to high-evaluated presentations as compared to low-evaluated presentations. Three significant clusters were observed in this contrast, which are indicated by blue arrows. Abbreviations: MOG, middle occipital gyrus; STG, superior temporal gyrus.

report by Downing et al. (2001). As a result, this ROI had a significant effect in the comparison of Hp vs. Lp conditions.

In summary, the results indicated the involvement of the visual areas especially associated with the perception of other people's actions in the distinct evaluation between Hp and Lp conditions. Observing an action would involve viewer engagement, which would lead to better evaluations in turn. To support this perspective, Experiment 3 analyzes the properties of the thin-sliced clips in terms of the magnitude of motion.

EXPERIMENT 3

We have shown that the evaluation of TED talks by their viewers were replicated by the thin-slicing technique (Experiment 1) and the brain regions associated with other people's actions are more active while watching the thin-sliced clips from highly rated talks (Experiment 2). Experiment 3 aimed to support the findings from Experiment 2 by quantifying the magnitude of motion in the thin-sliced video clip.

METHOD

Materials

Twenty thin-sliced videos of TED talks were analyzed. These videos were the same 30-sec thin-sliced videos as those used in the previous experiments. Therefore, they included 10 Hp and 10 Lp videos. The frame rate was 30 fps and the video size was adjusted to fit the 160×176 -pixel window.

Motion-magnitude analysis

The motion in these 20 videos was computed by optical flow (OF), which is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between a camera and a scene (Warren & Strelow, 1985). To compute OF, the widely accepted constraint is formulated as $I_x u + I_y v + I_t = 0$ (Horn & Schunck, 1981), where $I(x, y, t)$ is the intensity of a pixel and subscripts denote partial derivatives, and u and v denote the motions in x and y axes, respectively. Thus, the motion vector (u, v) at each location can be computed following the assumptions of brightness constancy and small movement. In this work, we applied an accurate OF method (Brox & Malik, 2011), and computed the motion vector of each pixel between frame t and frame $t + 3$. We skipped every two frames because those adjacent frames often had little change due to the high frame rate of 30 fps.

To consider the action in the thin-sliced TED talks, we summed up the magnitudes of motion vectors across one frame and then obtained the average of all n frames. This averaged sum of motion vectors was used as the quantitative index of motion magnitude for each video clip.

$$\overline{Sum} = \frac{1}{n} \sum_{t=1}^n Sum_t,$$

where the sum of motion vectors Sum_t for one frame Ω_t is formulated as

$$Sum_t = \sum_{(x,y) \in \Omega_t} \sqrt{u(x,y)^2 + v(x,y)^2}.$$

RESULTS AND DISCUSSION

Fig. 5A compares the motion magnitudes between Hp and Lp videos. As shown, the Hp videos showed greater motion magnitude than Lp videos, and the difference was statistically significant, $t(18) = 2.22$, $p = .047$, Cohen's $d = .95$ for log-transformed data.

Fig. 5. Motion-magnitude analysis. The sum of motion vectors was calculated for every three frames and averaged across 30 seconds. **A:** Grand averages of the averaged sum of motion vectors were computed separately for high-evaluated (Hp) and low-evaluated (Lp) presentations, and shown separately for the sets of videos used in Experiment 2. Error bars show SEM. **B:** The videos were sorted according to the averaged sum of motion vectors within Hp and Lp groups. Individual values for each video were then plotted according to the sorted order. For example, presentation clip 1 represents an Hp and Lp video with the highest motion magnitude within each group.

To examine the distribution of the magnitudes in more detail, we plotted the individual data for each video in Fig. 5B. This plot shows that, although overall Hp videos had greater magnitude of motion than did Lp videos, a large overlap was observed in the magnitudes between the two groups.

The significant difference in motion magnitude in Hp over Lp is consistent with the results of Experiment 2, which showed greater activation in EBA during Hp than Lp conditions. These results suggest an important role of the presenter's motion in presentation evaluations. However, it is notable that the distribution was largely overlapped. Considering this together with the fact that the current stimuli consisted of the top and bottom 10 videos from over 2,000 TED talks, the amount of motion alone is likely to contribute very little to the evaluation. The motion magnitude is more likely to influence the higher-order processing in the viewer's brain. This will be discussed in the following general discussion.

GENERAL DISCUSSION

This study examined the potential mechanisms underlying the evaluation of oral presentations. Experiment 1 applied the thin-slicing technique to oral presentations and showed that the rankings of the large video archive based on over 450 million ratings was represented by the rating of roughly 20 participants on 30-sec thin-sliced video clips. The motion magnitude of the clips coincided with the evaluation in Experiment 3. Together with the observation from Experiment 2 that motion-related visual areas were more active during Hp vs. Lp clips, presenter actions appeared to be one of the key factors for the

evaluation of oral presentations.

A possible model for the role of presenter actions may be discussed based on an interaction among three significant clusters found in the contrast between Hp and Lp in Experiment 2. As discussed in Experiment 2, the V2/V3d cluster is likely to be involved in global motion (Kujovic et al., 2013); therefore, the activation observed in this study was induced by the motion in the video. This motion-related activation would in turn influence the MOG-cluster activation, which is involved in the representation of other people's bodies through EBA (Downing et al., 2001). Thus, the difference in motion magnitude found in Experiment 3 would not merely represent the motion of each pixel, but would be more likely to be related more specifically to presenter actions.

EBA is known to be related to the observation of other people's actions as well as the orientation of attention by other people's gaze (Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006) and detection/evaluation of eye gaze in a social context (Kuzmanovic et al., 2009). These social signals could partly be affected by a top-down modulation from STG, which is another significant cluster in the current Hp vs. Lp contrast through its role in social cognition (Bigler et al., 2007; Pehrs et al., 2017).

The importance of nonverbal aspects has been reported repeatedly in the thin-slicing literature (Ambady & Rosenthal, 1992; E. Babad, 2005; Elisha Babad et al., 2003; Back, Schmukle, & Egloff, 2010). The current findings add further evidence in the context of oral presentations based on the combined BOLD response and video analysis, as well as the conventional questionnaire-based analysis. Particular emphasis was placed on an action, which is consistent with previous thin-slicing studies on teacher evaluations (Williams & Ceci, 1997). In addition, actions were also reported to play a critical role even in the evaluation of music performances (Tsay, 2013). The significant brain clusters reported here may be employed not only in the oral presentation, but also more generally in the evaluation of other people's performances.

In summary, the present results suggest that presenter actions play a critical role in the high evaluation of presentations. These actions are represented in the viewers' brain by MOG-cluster activation, including EBA, its top-down modulation from STG, and bottom-up modulation by V2/V3d. We should note that the present results were obtained from the presentations in a specific platform (i.e., TED). Further studies are needed to apply these achievements to fields such as education and coaching, which should use broader types of presentation and other method modalities. In addition to these future studies, the current study will open up new avenues for applying neuroscience techniques in not only advertising communications in the marketing field, but also for personal-level communications.

REFERENCES

- Ambady, N., Krabbenhoft, M. A., & Hogan, D. 2006. The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology*, **16**, 4–13.
- Ambady, N., & Rosenthal, R. 1992. Thin slices of expressive behavior as predictors of interpersonal consequences: A metaanalysis. *Psychological Bulletin*, **111**, 256–274.
- Ambady, N., & Rosenthal, R. 1993. Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, **64**, 431–441.

- Babad, E. 2005. Guessing teachers' differential treatment of high- and low-achievers from thin slices of their public lecturing behavior. *Journal of Nonverbal Behavior*, **29**, 125–134.
- Babad, E., Avni-Babad, D., & Rosenthal, R. 2003. Teachers' brief nonverbal behaviors in defined instructional situations can predict students' evaluations. *Journal of Educational Psychology*, **95**, 553–562.
- Back, M. D., Schmukle, S. C., & Egloff, B. 2010. Why are narcissists so charming at first sight? Decoding the narcissism-popularity link at zero acquaintance. *Journal of Personality and Social Psychology*, **98**, 132–145.
- Bigler, E. D., Mortensen, S., Neeley, E. S., Ozonoff, S., Krasny, L., Johnson, M., . . . Lainhart, J. E. 2007. Superior temporal gyrus, language function, and autism. *Developmental Neuropsychology*, **31**, 217–238.
- Bordier, C., Puja, F., & Macaluso, E. 2013. Sensory processing during viewing of cinematographic material: Computational modeling and functional neuroimaging. *NeuroImage*, **67**, 213–226.
- Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. 2009. Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, **43**, 703–706.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. 2004. Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, **86**, 599–614.
- Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. 2001. Brain areas sensitive to coherent visual motion. *Perception*, **30**, 61–72.
- Brainard, D. H. 1997. The psychophysics toolbox. *Spatial Vision*, **10**, 433–436.
- Brox, T., & Malik, J. 2011. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 500–513.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. 2001. A cortical area selective for visual processing of the human body. *Science*, **293**, 2470–2473.
- Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. 2006. Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage*, **33**, 406–413.
- Horn, B. K. P., & Schunck, B. G. 1981. Determining optical flow. *Artificial Intelligence*, **17**, 185–203.
- Kable, J. W., & Chatterjee, A. 2006. Specificity of action representations in the lateral occipitotemporal cortex. *Journal of Cognitive Neuroscience*, **18**, 1498–1517.
- Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., & Chatterjee, A. 2005. Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience*, **17**, 1855–1870.
- Kujovic, M., Zilles, K., Malikovic, A., Schleicher, A., Mohlberg, H., Rottschy, C., . . . Amunts, K. 2013. Cytoarchitectonic mapping of the human dorsal extrastriate cortex. *Brain Structure and Function*, **218**, 157–172.
- Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., & Vogeley, K. 2009. Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage*, **46**, 1154–1163.
- Murphy, H. A., Hildebrandt, H. W., & Thomas, J. P. 1997. *Effective business communications* (7th ed.). New York, NY: McGraw-Hill.
- Pehrs, C., Zaki, J., Schlottermeier, L. H., Jacobs, A. M., Kuchinke, L., & Koelsch, S. 2017. The temporal pole top-down modulates the ventral visual stream during social cognition. *Cerebral Cortex*, **27**, 777–792.
- Pelli, D. G. 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437–442.
- Pierno, A. C., Tubaldi, F., Turella, L., Grossi, P., Barachino, L., Gallo, P., & Castiello, U. 2009. Neurofunctional modulation of brain regions by the observation of pointing and grasping actions. *Cerebral Cortex*, **19**, 367–374.
- Tsay, C. J. 2013. Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 14580–14585.
- Walton, K. M., & Ingersoll, B. R. 2016. The utility of Thin Slice ratings for predicting language growth in children with autism spectrum disorder. *Autism*, **20**, 374–380.
- Warren, D. H., & Strelow, E. R. 1985. *Electronic spatial sensing for the blind: Contributions from perception, rehabilitation, and computer vision*. New York, NY: Springer.
- Williams, W. M., & Ceci, S. J. 1997. “How’m I doing?” Problems with student ratings of instructors and courses. *Change: The Magazine of Higher Learning*, **29**, 12–23.