Organizational Behavior and Human Decision Processes 124 (2014) 24–33

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

The vision heuristic: Judging music ensembles by sight alone

Chia-Jung Tsay

University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 20 July 2011
Accepted 8 October 2013
Available online 17 December 2013
Accepted by Richard Moreland

Keywords:
Social perception
Snap judgment
Judgment and decision making
Visual cues
Nonverbal behavior
Performance
Expertise
Groups
Team effectiveness
Cooperation

ABSTRACT

Team effectiveness and group performance are often defined by standards set by domain experts. Professional musicians consistently report that sound output is the most important standard for evaluating the quality of group performance in the domain of music. However, across six studies, visual information dominated rapid judgments of group performance. Participants (1062 experts and novices) were able to select the actual winners of live ensemble competitions and distinguish top-ranked orchestras from non-ranked orchestras based on 6-s silent video recordings yet were unable to do so from sound recordings or recordings with both video and sound. These findings suggest that judgments of group performance in the domain of music are driven at least in part by visual cues about group dynamics and leadership.

Introduction

Thirty-one classical musicians sit in silence on a stage, their instruments poised. The audience also sits expectantly, awaiting the conductor. Seconds later, the first notes of the symphony ring out – but the conductor has yet to appear. It seems far-fetched that a seemingly leader-less ensemble could begin playing at the same moment with such precision. Yet it is true: the world-renowned Orpheus Chamber Orchestra performs without a conductor (Lamb, 2001).

The success of such "unconducted" groups has often been attributed to a system of shared leadership (Hackman, 2002; Seifter & Economy, 2001; Traub, 1996). Such a system develops as team members influence each other and the team overall, harnessing their collective ability to create the conditions that foster team effectiveness (Hackman, 2005). At a more basic level, nonverbal and visual communication within unconducted groups facilitates coordination under dynamic conditions where creativity, spontaneity, and responsiveness (Thompson, 1967) are prized over more routine task parameters. Such in-process and unspoken mechanisms (Wittenbaum, Vaughan, & Stasser, 1998) can contribute to great performances by combining explicit coordination with more tacit coordination and mutual adjustment.

The astonishing phenomenon of the conductorless orchestra demonstrates vividly how subtle, visually based communication among group members can guide music ensembles to the creation of a coherent sound. Yet despite widespread recognition of coherent sound as the ultimate goal of top-performing music ensembles (Murnighan & Conlon, 1991), when it comes to the observation and evaluation of ensemble performance, visual information may dominate professional judgment. Recent research suggests that we overlook the degree to which visual cues can affect how we, as observers, judge the output of music ensembles: the sound of music (Tsay, 2013). For example, although both professional musicians and musical novices report that sound matters most to their judgment of music performance, they in fact rely primarily on visual cues when evaluating individual musicians (Tsay, 2013). In a set of experiments, Tsay found that both musical novices and experts identified the individual winners of live performance competitions through silent videos but were unable to do so through audio recordings or even recordings with both video and sound. This finding suggests that a striking visual dependence emerges even in a domain defined by auditory information.

In music competitions, a pianist's passion or a violinist's fluid and expansive gestures can sway a panel of judges. By contrast, we would expect the quality of an ensemble performance to be assessed based on much more than the idiosyncratic visual and affective information conveyed by individual performers. When multiple talented performers collaborate to make great music,

E-mail address: c.tsay@ucl.ac.uk

0749-5978
http://dx.doi.org/10.1016/j.obhdp.2013.10.003

Table 1
Summary of experiments.

Experiment	*N*	Stimulus type	Conditions	Versus at chance	Against other conditions
1	118	Professional group competition	V (video only of group), A (sound only of group), V/A (video plus sound of group)	V: 46.4%, $t(40) = 4.28$, $p < .001$ A: 25.8%, $t(33) = -2.71$, $p = .011$ V/A: 36.9%, $t(42) = 1.50$, p = n.s.	V vs. A: $t(73) = 4.90$, $p < .001$; Cohen's $d = 1.16$ V vs. V/A: $t(82) = 2.48$, $p = .015$ A vs. V/A: $t(75) = -3.05$, $p = .003$
2	130	Professional group competition	V (video only of group leader), A (sound only of group), V/A (video plus sound of group)	V: 43.8%, $t(50) = 4.90$, $p < .001$ A: 31.1%, $t(39) = -0.78$, p = n.s V/A: 28.9%, $t(40) = -1.80$, p = n.s.	V vs. A: $t(89) = 3.64$, $p < .001$ V vs. V/A: $t(90) = 4.59$, $p < .001$ A vs. V/A: $t(79) = 0.58$, p = n.s.
3	166	Professional group competition	V^1 (video only of group leader), V^2 (video only of group), V^3 (video only of non-leader)	V^1: 43.2%, $t(60) = 4.94$, $p < .001$ V^2: 47.8%, $t(59) = 5.22$, $p < .001$ V^3: 33.4%, $t(44) = 0.05$, p = n.s.	V^1 vs. V^2: $t(119) = -1.36$, p = n.s. V^1 vs. V^3: $t(104) = 3.11$, $p = .002$ V^2 vs. V^3: $t(103) = 3.74$, $p < .001$
4	283	Professional group competition	V^1 (video only of group leader), V^2 (video only of group), A (sound only of group), V/A (video plus sound of group)	V^1: 41.4%, $t(72) = 3.84$, $p < .001$ V^2: 55.3%, $t(52) = 9.60$, $p < .001$ A: 26.5%, $t(66) = -3.86$, $p < .001$ V/A: 36.8%, $t(70) = 1.70$, p = n.s.	V^1 vs. V^2: $t(124) = -4.44$, $p < .001$ V^1 vs. A: $t(138) = 5.38$, $p < .001$ V^2 vs. A: $t(118) = 10.10$, $p < .001$ V^1 vs. V/A: $t(142) = 1.56$, p = n.s. V^2 vs. V/A: $t(122) = 6.01$, $p < .001$ A vs. V/A: $t(136) = -3.80$, $p < .001$
5	172	Professional orchestras	V (video only of group), A (sound only of group), V/A (video plus sound of group)	V: 64.3%, $t(61) = 8.13$, $p < .001$ A: 53.0%, $t(55) = 1.30$, p = n.s. V/A: 60.6%, $t(53) = 5.10$, $p < .001$	V vs. A: $t(116) = 3.90$, $p < .001$; Cohen's $d = 0.72$ V vs. V/A: $t(114) = 1.37$, p = n.s. A vs. V/A: $t(108) = -2.41$, $p = .017$
6	193	Professional group competition	V^1 (video only of group leader), V^2 (video only of group), A (sound only of group), V/A (video plus sound of group)	V^1: 35.8%, $t(46) = 1.09$, p = n.s. V^2: 40.1%, $t(37) = 2.30$, $p = .027$ A: 32.4%, $t(53) = -0.48$, p = n.s. V/A: 33.7%, $t(52) = 0.18$, p = n.s.	V^2 vs. A: $t(90) = 2.26$, $p = .026$ V^2 vs. V/A: $t(89) = 1.78$, $p = .079$ All other comparisons, p = n.s.

group dynamics should have a strong impact on the overall process and performance.

When evaluating the performance of musical groups, both novices and professional musicians report that their judgments are based upon the overall *sound* the musicians produce. For example, in interviews of British string quartets, Murnighan and Conlon (1991) found that the collective task of chamber music ensembles is "to reach a high level of coordinated sound". The literature on team effectiveness (Hackman, 1987) would also support the notion that sound should be taken as the gold standard for evaluation; after all, professional ensemble musicians themselves deem the production of "transcendent, glorious *sound*" as their goal (Murnighan & Conlon, 1991, p. 167), suggesting that sound is most important to their evaluation of music ensembles.

This paper explores the degree to which visual information influences expert judgments of group performance. A set of six experiments considered the degree to which visual information allows quick estimates of the outcomes of international ensemble competitions and professional rankings of symphony orchestras. An assessment of the relative contribution of visual vs. auditory information in the domain of music allows for the most conservative test of the primacy of visual cues.

There are several ways in which this work extends research and theory, with important practical implications. First, building on recent research, the current studies serve as the first empirical investigations in support the notion of the *vision heuristic*, which describes the way in which people use visual information more than they are aware of, more than they rely on auditory information, and beyond what they would endorse or choose with greater reflection. Whereas the earlier work focused on perceptions of individual performance (Tsay, 2013), the present research focuses on judgments of work groups and teams, group processes, and team performance. Second, this research introduces the thin-slices phenomenon to perceptions and outcomes of group interactions. Third, this research explores ways in which the standards and values of professionals are at odds with how they actually evaluate group output. Fourth, in a continued investigation of professional standards for the judgment of team effectiveness, this research examines the degree to which the vision heuristic can transcend domain knowledge, experience, and expertise. Finally, this work offers an investigation of the visual cues underlying perceived status, leadership, and group dynamics, and the influence of these factors on professional judgment.

Thin-slices research

Key decision-makers are more likely to have informal and spontaneous interactions with others than the managers of earlier generations were (Mintzberg, 1975). In various arenas of assessment, we have become more dependent on rapid social judgment, or the impressions and evaluations formed on the basis of minimal verbal and nonverbal cues, which contribute to more enduring perceptions. These initial impressions may affect our assumptions about others, which can then fundamentally change our own behaviors and the attitudes and behaviors of our interaction partners (Word, Zanna, & Cooper, 1974) and ultimately affect more general individual and organizational outcomes.

"Thin slices" of nonverbal behavior have been shown to have a strong impact on social judgment in a wide range of areas, including education, medicine, and personality assessment (Ambady & Rosenthal, 1993). That body of work suggests that we evaluate others quickly and automatically, such that impressions made in a few seconds can be highly predictive of impressions made after much longer periods of time. Such impressions also reveal other important information, such as internal state and moods, personality traits, and social and interpersonal relationships (Ambady, Bernieri, & Richeson, 2000; Ambady, Conner, & Hallahan, 1999).

Making judgments on the basis of thin slices requires interpretation of nonverbal and visual cues, which become the basis of our interpretation of future interactions. Previous research shows significant correlations between evaluations based on thin slices and more long-term evaluations of interest to organizational life, such as job performance and employment interviews (Ambady et al., 2000). More recent research points to an association between facial characteristics and consequential decisions and outcomes,

such as candidates' election results (Todorov, Mandisodza, Goren, & Hall, 2005), and even the impact of CEOs on the financial performance of their companies (Rule & Ambady, 2008, 2011).

Beyond perceptions of others, recent research has begun to unveil how thin slices of conversational dynamics, including visual information, predict dyadic outcomes in negotiations (Curhan & Pentland, 2007). Using microcoding, this research demonstrated that activity level, conversational engagement, variation in speech pitch and volume, and vocal mirroring that occur within the first 5 min of negotiations were highly predictive of negotiated outcomes. Similarly, microcoding of affect during the first minutes of marital conflict was found to be predictive of marital outcomes over several years (Carrere & Gottman, 1999).

To further our understanding of the broader relevance of thin slices for different levels of analysis, it is important to conduct a test of the primacy of visual information in judgments of group processes and performance. Because of the consistent belief that auditory cues are the main basis for judgment in the domain of music, this domain offers an extreme test of the importance of visual cues for evaluating group performance. Furthermore, team and organizational effectiveness is in part defined by the standards of "interested stakeholders," or the evaluation of output by legitimate reviewers (Hackman, 1987). Through settings such as international chamber ensemble competitions, participant judgments allow the extrapolation of the professional evaluation of quality.

The impact of visual vs. auditory information

Although people often rely on visual cues to understand their environment, it is unclear whether this reliance results in optimal decisions and outcomes. Some work has suggested that the influence of visual information can be detrimental to our evaluation of others and to the accuracy of social judgment (Olivola & Todorov, 2010). For example, visible demographic cues can lead to us to use negative stereotypes when categorizing others (Allport, 1954; Dovidio, Evans, & Tyler, 1986; Duncan, 1976; Gaertner & McLaughlin, 1983; Gilbert & Hixon, 1991; Sager & Schofield, 1980; Smith, 1984; Word et al., 1974). In contrast, recent work suggests that certain nonverbal cues may be more relevant than previously assumed in our predictions of others' behaviors. If gestural displays enact psychological, physiological, and behavioral changes (Carney, Cuddy, & Yap, 2010), and these changes result in differences in task performance, then these nonverbal cues may actually hold predictive value for the assessment of the quality of performance.

It is thus important to explore when and to what extent visual cues are used as the primary source of information, and the relative contribution of visual information vs. other types of information, in our judgments and decisions. Literature from a range of disciplines compares the processing of auditory vs. visual information (Bryce & Yalch, 1993), such as the disruptions and synergies among different channels of communication. The role that the two types of information play in interpersonal communication, for example, is complex and nuanced with regard to the conditions under which sight vs. sound will dominate (Amabile & Kabat, 1982; DePaulo, Lassiter, & Stone, 1982; Harrigan, Wilson, & Rosenthal, 2004; Tanaka et al., 2010). Still, the presence of multiple channels of information can aid learning because different types of information, as part of a dual-coding system, are hypothesized to not compete with each other (Paivio, 1978).

Different channels of information may also unexpectedly shift our perceptions of seemingly unrelated modalities. For example, the McGurk effect is one instance in which visual input can fundamentally change how we perceive the sound of words (McGurk & MacDonald, 1976). If an audio recording of the phoneme "ba" were combined with the visual display of "ga," people would hear "da." Such an interaction between visual and auditory cues cannot be suppressed even by knowledge of the effect (McGurk & MacDonald, 1976).

Multi-modal interactions have captured the attention of other researchers as well, suggesting intriguing sensory effects on presumably unrelated outcomes of interest (Ackerman, Nocera, & Bargh, 2010; Jousmäki & Hari, 1998), such as the influence of auditory information (Zampini & Spence, 2004) and visual information (Hoegg & Alba, 2007) on taste. The work on verbal overshadowing also points to the many ways in which verbalization disrupts the processing of nonverbal information (Lane & Schooler, 2004), such as those associated with face identification (Schooler & Engstler-Schooler, 1990), memory for taste (Melcher & Schooler, 1996), and descriptions of visual stimuli (Brandimonte, Schooler, & Gabbino, 1997).

Of particular interest are domains in which non-visual information is valued and explicitly cited as central to decisions about performance. When it comes to politicians and actors, we might intuitively expect visual cues to account for a large proportion of the variance in evaluation. However, there is little evidence of how strongly visual information is weighted when there is consensus that non-visual information should form the core of evaluation. Music, a domain consistently regarded as auditory in nature, would provide a strong test of the primacy of visual information (Paivio, 1978).

Yet even in this domain, recent research suggests that visual cues are strongly weighted and that this tendency may operate at a nonconscious level. Professional musicians cite a lack of confidence (Burson, Larrick, & Klayman, 2006) in their judgments of music performance when they have no access to sound, unaware that they can approximate expert decisions through silent videos (Tsay, 2013). People have little conscious insight into their cognition (Bargh, 1996, 1997, 2006; Greenwald, McGhee, & Schwartz, 1998; Nisbett & Bellows, 1977; Nisbett & Valins, 1971; Nisbett & Wilson, 1977), dissociating between the type of information they report valuing and the type of information they actually use.

Because information that is highly available can dominate predictions (Tversky & Kahneman, 1973), it may be that visual cues are simply more available than other types of sensory information and thus literally most visible. The primacy of visual cues may emerge because of their greater salience over other modes of information, such as sound. One of the main critiques of rational models of behavior and choice is that decision makers often only have access to incomplete information (Kahneman, Slovic, & Tversky, 1982). However, even when more complete information is available, we may still rely on incomplete but more salient information.

Research has also pointed to the conditions under which people process peripheral more than central cues. For example, heuristics-based processing may result when there are constraints on someone's ability or motivation to process information systematically (Malhotra & Bazerman, 2008; Petty & Cacioppo, 1986), as when cognitive load is high or when an issue seems unimportant. Other work has demonstrated the many ways in which cognition can be affected by irrelevant or nondiagnostic information (Castellan, 1973), such as when evaluations of outcome quality are affected by information about preparation time (Chinander & Schweitzer, 2003), reaction times to stimuli are altered by the location of nondiagnostic information (Kreuger, 1973; Simon & Rudell, 1967; Simon & Small, 1969), or the predictive power of diagnostic information is diluted in the presence of nondiagnostic information (Nisbett, Zukier, & Lemley, 1981).

In the domain of music, visual cues are considered peripheral to the assessment of the quality of performance, despite their possible role in contributing to more effective group dynamics and coordination in ensembles. Nonetheless, visual cues may be more salient to judges and audiences. The present research explores the ways in which experts overweight visual cues in their judgments of a group's performance relative to auditory output, the purported main substance of their evaluations.

In this paper, six experiments examine how visual cues come to affect judgments of the quality of music performance by groups. These experiments include comparisons between visual and auditory cues in an effort to see which cues have the greatest impact on professional judgments of musical group processes and performance.

Experiment 1: A first test of the primacy of visual cues for musical groups

Experiment 1 offered a first test of the dominance of visual information in assessments of music performance by groups. Live international chamber ensemble competitions offered a unique opportunity for the extrapolation of professional judgment by comparing judgments of thin slices of group performance made by novice study participants against the actual competition outcomes. These prestigious competitions serve as critical career milestones designed to launch professional careers through international publicity, concert tours, and financial support. Because of the consequential nature of these top competitions, organizations invite professional musicians to serve as judges. These musicians usually have both expertise and experience, often having belonged to renowned, prize-winning chamber ensembles themselves.

Method

Participants

A total of 118 participants (ages 18–50; 58 males, 59 females[1]) with little to no training in classical music volunteered for the experiment. Participants were asked to categorize themselves into one of four levels of training: none, childhood/amateur, conservatory/pre-professional, or professional. They were recruited from a community sample in the Northeastern United States and were paid $20 for their participation in an hour-long set of unrelated studies that included the current experiment.

Materials

Stimuli were excerpted from publicly available recordings of international ensemble competitions, including the Fischoff Competition, the Saint Paul String Quartet Competition, and the Terem International Music Competition. Because the finalist groups performed different compositions, 6-s excerpts were selected that showcased approximately equal technical difficulty, volume of sound, and musical intensity. The length of 6 s was chosen based on the previous literature using the thin-slices paradigm (Ambady & Rosenthal, 1993) and held predictive power for evaluations of much longer periods of time. Earlier experiments (Tsay, 2013) also showed there were no significant differences between patterns that emerged through 6-s and 60-s clips.

All excerpts were pretested on a separate sample of 15 professional musicians with an average of 17.93 years of formal training. Repeated-measures ANOVAs showed that there were no significant differences across the excerpts on the dimensions just listed (all p's > .05). Throughout all the experiments, stimuli were presented in random order at multiple levels: the competitions appeared in random order, and within each competition set, the three excerpts also appeared in random order.

Procedure

The three top finalist groups in each of eight ensemble competitions were presented to participants in random order. In a between-subjects design, participants received either sound-only, video-only, or video-plus-sound recordings of the same excerpts of each group performance. Participants were run both in groups and as individuals. In the group settings, each participant had access to his or her own headset and monitor and was separated from other participants through partitions. This ensured that participants did not see or hear the responses of other participants.

Based on the set of three finalist chamber ensembles, participants were then asked to select the winning group in each of the competitions. Participants were told that their choices would contribute to a better understanding of decision making in professional domains. No participants reported visual recognition of any of the chamber ensembles; thus, any differences in rates of identifying the winning groups in the video-only condition could not be attributed to existing knowledge about the groups or the competition. If this selection were done randomly among the three choices in each competition, then the rate of identification of the winning group would have been at chance (33%). Finally, in a forced-choice item, participants were asked to report whether sound, visual cues, or other types of information mattered most in their evaluation of ensemble music performance. Participants also had the opportunity to explain how they chose the winning groups and what types of cues they relied on for their decisions, and to provide any additional feedback regarding the experiment.

Results and discussion

As expected, a majority of participants (80.7%) identified sound as the most critical information for the evaluation of group performance, $\chi^2(2, N = 119) = 122.13$, $p < .001$. Yet with video-only recordings, participants were significantly better than chance (46.4%) at identifying the winning groups, $t(40) = 4.28$, $p < .001$. They were actually below chance (25.8%) with sound-only recordings, $t(33) = -2.71$, $p = .011$. A separate item analysis was performed, averaging across participants and comparing each item or specific competition trial average across all participants in one condition vs. the average across all participants in another condition, to test whether the effects held across trials. The analysis indicated that these effects were robust across all eight competitions, $t_2(14) = 4.24$, $p = .001$. Finally, when provided with recordings with both sound and video, participants operated at chance levels (36.9%), $t(42) = 1.50$, p = n.s. (see Table 1)

Musical novices were thus able to quickly identify the winning groups of top international ensemble competitions through visual cues but not through the sound of the performances. Previous research demonstrated that for short recordings in all three conditions (video-only, sound-only, video-plus-sound), there are no significant differences between novices and professional musicians in their capacity to identify the actual winners of classical music competitions (Tsay, 2013).

In this experiment, the ensembles were similar in size within each competition; thus, the size of the groups was unlikely to have influenced the results. Because the competitions all had a similar age-range restriction, there was little difference in the overall age of the groups. Similarly, in most competition trials, there were at least two groups that had the same gender and race ratio, thus making it unlikely that these salient demographic characteristics significantly contributed to differences in judgments or attention.

[1] Participants who did not report their gender were not included in the calculation.

Participants who received video-plus-sound recordings were also no better than chance at identifying the competition winners. These findings suggest that the presence of sound actually distracted people from the actual outcomes, providing support for the extrapolation that the original professional judgment was based primarily on visual rather than auditory cues. For example, participants cited visual communication, expression, movement, passion, involvement, and eye contact as reasons for choosing the groups that they thought were competition winners.

Experiment 2: Perceptions of leadership through visual cues

Experiment 2 investigated which visual cues most influence decision making about musical group performance. The first experiment suggested that even in the presence of more information, such as the video-plus-sound recordings provided, the auditory information was not fully used. Recent work points to the role of visible passion and movement in leading professionals to assumptions about the quality of performance. For example, in the previous experiments (Tsay, 2013), visual information allowed participants to distinguish between winners and non-winners on criteria such as passion, yet sound recordings did not allow them to do so. In Experiment 1, the original expert judges may have been particularly influenced by visually salient individuals when they should have been attuned to the dynamics of the ensemble. If there is one individual who has the most influence over musical decisions (Murnighan & Conlon, 1991), who would be particularly expressive for purposes of coordinating the ensemble, and who thus would be most salient to observers, it would be a group's leader. This position is well-established for music ensembles and is the basis around which relationships within ensembles are built and negotiated (Murnighan & Conlon, 1991).

Using the same stimuli as in Experiment 1 and another between-subjects design, 130 novice participants (63 males, 67 females) were presented with either sound-only recordings of the groups, video-plus-sound recordings of the groups, or video-only recordings of the *group leaders only*. For example, in string quartets, the group leader is the first violinist; in piano trios, the violinist serves as the leader. Experiment 2 explored whether the presence of visual information about a group's leader allows participants to identify the outcomes of ensemble competitions.

Replicating the findings from Experiment 1, only 31.1% of those who received sound-only group recordings identified the winning groups, $t(39) = -0.78$, p = n.s., and only 28.9% of those who received video-plus-sound group recordings did so, $t(40) = -1.80$, p = n.s. Thus, for both types of recordings, participants performed at about chance (33%). In contrast, when observing silent video recordings featuring only the group leaders, 43.8% of participants selected the winning groups—a rate significantly above chance, $t(50) = 4.90$, $p < .001$, significantly above those with sound-only group recordings, $t(89) = 3.64$, $p < .001$, and significantly above even those with video-plus-sound group recordings, $t(90) = 4.59$, $p < .001$. Also replicating the results of Experiment 1, a significant majority of participants (89.2%) again said that sound was the most important factor in their evaluations of group performance, $\chi^2(2, N = 130) = 184.45$, $p < .001$.

Experiment 2 offered stronger support for the vision heuristic. Providing just a subset of visual information—the silent view of a single individual—allowed participants to identify the actual outcomes of ensemble competitions, and at a rate greater than that of participants with audio or live performance recordings of the entire groups. These results suggest that the original expert judges were highly influenced by visual information about the group leaders, even when they reported that the ensemble sound was the basis for their evaluations.

Experiment 3: A visual comparison of leaders vs. non-leaders

The findings of Experiment 2 suggested that visual information about just one individual, a group's leader, influenced the professional judgment of group performance above and beyond the impact of the sound of the group. Even though visual cues about a group's leader seem to provide less basis for judgment than information about the entire group, visual cues such as movement, focus, control, and expression still allowed participants to identify actual competition outcomes at rates significantly above chance. This was in contrast to how participants performed with the sound-only and video-plus-sound group recordings: no better than chance.

On the one hand, information about a group's leader could convey a significant amount of information about the overall group performance because he or she has the primary responsibility for coordinating the performance and tends to be representative of the average quality or competence of the group. On the other hand, it may be that information about any group member conveys enough information for participants to make informed evaluations about the group overall.

In Experiment 3, 166 novice participants (69 males, 94 females) were presented with one of three types of recordings: video-only recordings of the groups, video-only recordings of group leaders, or video-only recordings of group members who were not leaders. The non-leaders were excerpted from the original recordings by selecting an individual standing on the opposite side of the stage from the leader. The experiment implemented a between-subjects design using the same stimuli as in the previous two experiments.

With the video-only group recordings, 47.8% of participants were able to select the actual winners, at rates significantly above chance, $t(59) = 5.22$, $p < .001$. Through video-only recordings of the group leaders, 43.2% identified the winners, at rates significantly above chance, $t(60) = 4.94$, $p < .001$. Yet when visual cues were provided about the group members who were not leaders, only 33.4% of participants identified the actual winners, a rate not significantly better than chance, $t(44) = 0.05$, p = n.s. Visual cues associated with group leaders were significantly more informative for judges in their evaluation of group performance as compared to cues associated with non-leaders, $t(104) = 3.11$, $p = .002$.

The findings from Experiment 3 suggest the importance of leadership cues. Novice participants were particularly influenced by the leaders of chamber ensembles in their evaluations of group performance, given that decisions made based on information from group leaders alone allowed participants to better approximate the original professional decisions about the groups than did information about non-leader group members.

Experiment 4: The relative contribution of visual leadership and group dynamics

To disentangle the relative contributions of visual information about leadership vs. group dynamics, Experiment 4 pitted silent video recordings of the group leader alone against silent video recordings of the entire group. Again using a between-subjects design and the same stimuli as in the earlier three experiments, 283 novice participants were presented (130 males, 127 females[1]) with one of four types of recordings: sound-only recordings of the groups, video-plus-sound recordings of the groups, video-only recordings of the groups, or video-only recordings of only the group leaders. Given that the participants in this study were novices, they would not be aware of which position or instrument would have been designated the group leader, and thus unaware of such leadership roles when they were randomly assigned to the condition in which they saw only the group leaders.

Although the group leader may be most representative of the quality of the musicians in an ensemble, the inclusion of additional information about the entire group is likely to reveal not only information about the leader but also relevant information about how well the group functions. Despite the importance of the leader's role, both the production of musical content and real-time performance dynamics can depend on other group members.

Replicating the previous results, with sound-only group recordings, only 26.5% of participants were able to select the winning groups, a rate lower than chance, $t(66) = -3.86$, $p < .001$. Similarly, with video-plus-sound group recordings, only 36.8% of participants selected the winners, $t(70) = 1.70$, p = n.s. However, with video-only recordings of the group leaders, 41.4% of participants identified the winners, a rate significantly above chance, $t(72) = 3.84$, $p < .001$. With video-only recordings of the entire groups, 55.3% of the participants (again, a rate above chance) were able to identify the winners, $t(52) = 9.60$, $p < .001$. Having access to visual cues about the entire group was even more informative for judgments of group music performance as compared to having access to information about just the group leader, $t(124) = 4.44$, $p < .001$.

These findings again suggest that the visual components of ensemble music performances influence judgments of the music, despite the consistent recognition of sound as central to the evaluation of performance. Visual information about the groups and group leaders allowed participants to arrive at the professionally determined competition outcomes at rates significantly greater than chance; the previous studies showed that this was not the case when sound recordings were provided, despite participants' beliefs about the greater importance of ensemble sound. Visually conveyed cues about leadership and group dynamics, rather than the sound output, appeared to have influenced the original judges' decisions about the quality of group performance, even in the domain of music.

Experiment 5: A test of the primacy of visual cues in large groups

The previous four experiments suggest that visible cues about group leaders and group dynamics contributed to the professional judgment of group performance, whereas ensemble sound had little influence on competition outcomes. Experiment 5 posed a test of the primacy of visual cues in symphony orchestras. The fact that the designated leaders were not visible in these recordings and the group size made shared leadership unlikely (Khodyakov, 2007) allowed for a closer investigation of the influence of group dynamics on the judgment of performance.

Recent research has demonstrated the role that physical synchrony (Valdesolo, Ouyang, & DeSteno, 2010) and emotion contagion (Barsade, 2002) can play in facilitating social cohesion, coordination, and the ability to pursue joint goals. If physical synchrony and shared affect can facilitate better performance in the form of higher-quality auditory output, this suggests that these visual cues are informative for the judgment of group performance, as they reflect the quality of performance.

Method

Participants

A total of 172 participants (ages 18–50, 83 males and 87 females) volunteered for the experiment.[1] Participants had little to no training in classical music.

Materials

Because no competitions exist for symphonic performance at the international level, professional rankings of the top orchestras in the world were used. Six-second recordings of live performances were selected from the top ten orchestras in the world, including the Royal Concertgebouw Orchestra (Amsterdam), the Berlin Philharmonic, the Vienna Philharmonic, the London Symphony Orchestra, the Chicago Symphony Orchestra, the Bavarian Radio Symphony, the Cleveland Orchestra, the Los Angeles Philharmonic, the Budapest Festival Orchestra, and the Dresden Staatskapelle. The recordings were then matched with identical excerpts of the same repertoire performed by non-ranked orchestras, such as regional or university-based groups. Although the musical content performed was identical in each pair, the selection of regional or university-based groups ensured that, based on professional judgment and the different processes involved for placement into the respective groups, there should be significant differences in the overall quality of the ensembles within each pairing. All groups performed in comparable attire and at comparable concert venues.

Procedure

Ten pairs of orchestral performances were presented to participants. In a between-subjects design, participants received either sound-only, video-only, or video-plus-sound recordings. They were then asked to select the top-ranked group out of each of the pairs of symphony orchestras. Selection at chance would have resulted in a rate of 50%. Finally, participants were asked to report whether sound, visual cues, or other cues mattered most in their evaluation of orchestral music performance. The comparison of novice evaluations with the expert judgment of symphony orchestras worldwide allowed for another extrapolation about the process of professional judgment in the domain of music.

Results and discussion

With video-only recordings, participants were significantly better than chance (64.3%) at identifying the top-ranked orchestras, $t(61) = 8.13$, $p < .001$. With sound-only recordings, they performed at chance (53.0%), $t(55) = 1.30$, p = n.s. Participants performed significantly better with video-only recordings than sound-only recordings, $t(116) = 3.90$, $p < .001$; Cohen's $d = 0.72$. When provided with recordings with both sound and video, participants were also above chance (60.6%), $t(53) = 5.10$, $p < .001$, performing better than those who received sound-only recordings, $t_1(108) = 2.41$, $p = .017$.

In the evaluation of larger groups such as symphony orchestras, the primacy of visual information again emerged. Novice participants who were provided with video-only recordings were able to approximate expert judgment at levels significantly above chance. These findings included a slight departure from the earlier results and may be due to the reduced capacity of musical novices to use audio information. This experiment included only identical repertoire excerpts, which reduced the more discernable variance in the quality of audio content for those with little training in music. In this case, the addition of audio content in the video-plus-sound recordings would not add useful information for novices. Even so, those who received video-only recordings were still best able to approach the original expert evaluations.

Experiment 6: A test of musical expertise

The earlier experiments suggested that visual information about group leaders and group dynamics allows even musical novices to arrive at the decisions made by professional musicians after lengthy expert evaluations of live performances. Although this suggests that professional musicians overweight visual information, it may also be that the previous novice participants did not have the training or knowledge needed to be able to appropriately use auditory information. Thus, novices may have had to depend on visual

information because it was the only type of information they could use to discern quality of performance.

Experiment 6 tested whether musicians with professional training and experience would also depend heavily on visual information in the judgment of music ensembles. Using the same between-subjects design as in Experiment 4, a sample of 193 professional musicians (110 males, 83 females) was presented with one of four types of recordings: sound-only recordings of the groups, video-plus-sound recordings of the groups, video-only recordings of the groups, or video-only recordings of only the group leaders. On average, these musicians had 14.85 years of training in classical music, and all identified themselves as having conservatory training and/or professional experience in the domain of music.

As before, experts (81.3%) again identified sound as the most important piece of information in the evaluation of group performance, $\chi^2(2, N = 193) = 200.72$, $p < .001$. Yet despite their beliefs about the importance of sound, and despite their presumably greater ability to use sound to evaluate musical output, only 32.4% of them were able to select the winning groups using sound-only group recordings. Although this was a higher rate than in the previous experiments, with novice participants, it was still not significantly different from chance, $t(53) = -0.48$, p = n.s. Even with video-plus-sound group recordings, only 33.7% of the participants selected the winning groups, $t(52) = 0.18$, p = n.s. With video-only recordings of the group leaders, 35.8% of participants identified the winning groups, $t(46) = 1.09$, p = n.s. Finally, with video-only recordings of the entire groups, 40.1% were able to identify the winning groups, $t(37) = 2.30$, $p = .027$.

The chi-square test of independence suggests that there is a relationship between the condition and the rates of correspondence between participant decisions and actual outcomes, $\chi^2(4, N = 1{,}343) = 14.76$, $p = .005$. To determine which cell or cells produced the statistically significant difference, the standardized residuals were examined. The only standardized residual that emerged as larger ($z = 2.5$) than the critical value (alpha of 0.05, ±1.96) was for the cell that suggests that more participants in the video-only condition selected the first-place group as the winner than would have been expected.

Although these experts were more likely to note their lack of confidence when only a limited set of visual information was available to them, and thus may have tempered their use of such information in the latter two conditions, such information still led them to better approximate the original outcomes at levels that auditory information did not.

The results from Experiment 6 present more compelling evidence pointing toward the vision heuristic, given that both novices and experts alike depended on visual information in the evaluation of ensemble music performance. There were no significant effects of gender, age, or years of music training on whether participants' selections were more or less likely to match the actual competition outcomes. Future work may investigate whether there are other individual differences that would predict a reliance on the vision heuristic, such as musical, bodily-kinesthetic, or interpersonal intelligence (Gardner, 1993).

Although professional musicians were just as likely to acknowledge the importance of sound to their judgments in this domain, they again appeared to be unaware of the extent to which they relied on visual cues about group dynamics when assessing the quality of music performance. If novices spontaneously use visual cues to assess the value of others, perhaps professionals are just as or even more likely to do so, and indeed also feel entitled to do so, because of their knowledge of their own expertise (Leyens, Yzerbyt, & Schadron, 1992; Yzerbyt, Schadron, Leyens, & Rocher, 1994). Such findings highlight the ways in which expertise and knowledge in a domain may not temper a more basic heuristic, one that guides experts away from the information that they have been trained to use and on which they place the greatest value.

General discussion

Whether the performance involves one musician, a chamber ensemble, or a symphony orchestra, people appear to overweight visual information in their evaluation of music performances. This vision heuristic rapidly guides people, even domain experts, toward judgments. Six experiments suggest that music ensembles that garner international professional recognition may be those that convey more convincing visible leadership and group dynamics.

Chamber ensembles and professional musicians may claim that "the ultimate quartet [...] astounds its members and its listeners" (Murnighan & Conlon, 1991; p. 170). Yet the mere presence of sound in the video-plus-sound recordings actually detracted from the predictive power of video-only recordings. Furthermore, with sound-only recordings, participants were significantly less able to identify the winning groups. This research suggests that the ultimate music ensemble astounds not its listeners, but its *viewers*.

Despite the discrepancy between what experts value and the information they appear to actually use when evaluating musical groups, the nonconscious reliance on visual information may still hold some relevance even in this domain. Given that the production of music is necessarily mediated by physical behavior, and given that this is particularly the case when ensembles and group coordination are involved, visual information should contribute to and be predictive of outcomes. Work on jazz ensembles has discussed the value of nonverbal signals, such as synchronized swing, exchanged by collaborative musicians during performances (Gloor, Oster, & Fischbach, 2011). It may be that the thin-slice measure of visual information can be just as relevant as, or even more relevant than, the auditory musical output itself. Despite what domain experts believe about the value of auditory output, various important qualities conveyed by visual information may be more reliably sampled and evaluated.

Perceptions of "groupiness" or entitativity (Campbell, 1958) may also help account for the apparent importance of visual cues in evaluations of music performance. Visual cues involving such matters as similarity, proximity, and shared goals or outcomes among the members of a musical ensemble can serve as relevant information about the quality of the group's performance; ensemble training often reflects these beliefs, as seen in coordinated themes such as the performance attire or stage entrance of musicians. It may be that our impressions of entitative groups become spontaneous and organized (Hamilton & Sherman, 1996; Hamilton, Sherman, & Maddox, 1999; Sherman, Castelli, & Hamilton, 2002) around more coherent themes or traits. If the coordinated action or entitativity in top-performing groups is high, this may explain why visual characteristics are so persuasive to judges and participants, for whom coordinated movement is a strong signal of differences among ensembles.

Finally, the ways in which people are often unable to describe the reasons underlying their attitudes and judgments (Wilson & Schooler, 1991) may provide some perspective on the findings in this paper. When people consider the information they believe matters most, they appear to focus on criteria that are less relevant to the factors that truly matter. Indeed, those who listened to audio stimuli made ratings that diverged farther from expert evaluations than did those who viewed the ensembles without sound.

For both scholars and practitioners, it remains important to reflect on the actual value of visual information for evaluations of both individuals and groups. Our unwillingness to report such reliance suggests that we believe that sound is at the heart of what we

should value in the domain of music because sound is more likely to produce wiser decisions. This notion motivated the development of "blind" auditions and screenings (Goldin & Rouse, 2000) to determine with greater objectivity which musicians would be most fit for coveted orchestral positions. Yet the present findings suggest that what is chosen under conditions meant to foster impartiality may be inconsistent with the quality of performance later perceived and experienced by expert judges and lay audiences.

There may be a critical and unrecognized discrepancy between what is conveyed through sound vs. vision; what is understood as a great performance through radio stations and MP3 recordings may be very different from what is deemed as exceptional during live concerts. My research points to implications about the validity of different types of performance evaluation, as different outcomes may result based on which type of information is available and evaluated. Although audio recordings may lead people to prefer one musician, live performances may lead them to prefer another musician. If live performances entail the overweighting of visual information and the neglect of sound itself, then it becomes necessary to question the role of blind auditions in professional selection processes, because they can contribute to potential "mispredictions and mischoices" (Hsee & Zhang, 2004).

Qualitative data regarding participants' reasoning about their choices may shed further light, through content analyses, on the particular types of cues most associated with quality of performance, as well as levels of confidence about choices. Although participants were unaware of how much they relied on visual cues when evaluating the performances (given their explicit consensus on sound as being the most important source of information), the experiments provided free-response items that allowed participants to describe what types of visual information they prioritized. These descriptions may contribute to a better understanding of which cues allow for greater consensus and professional consistency in preference.

Questions of external validity, and of which information channel leads to better outcomes, are also worthy of investigation. These questions may be examined by following up on the short- and long-term outcomes of those selected as "winners" in the sound-only, video-only, and video-plus-sound conditions. Such follow-up will help determine whether certain types of information are more predictive of the quality of performance over the course of a career. Future work can explore whether expert judges' evaluations are predictive of short-term and/or long-term success and, accordingly, which type of information should be most valued.

If "winners" selected under the sound-only condition are more successful in the long run, then that would raise strong concerns for the state of training programs and selection policies. If domain expertise cannot overcome a more basic dependence on visual information, and if we are better off not having visual information at all, then it would be critical to explore how to help professionals come to a fuller understanding of the information that they use in their roles as decision makers. It would also be important to reexamine how professional organizations routinely standardize and institutionalize decision processes regarding identifying, promoting, and rewarding talent.

Although it may not be possible to eliminate our visual dependence altogether, we may be able to attenuate its effects on judgment. For example, ongoing research suggests that reducing the cognitive demands on decision makers may reduce the likelihood of relying on the vision heuristic. During music competitions, judges are often confronted with hours of repertoire per contestant that require them to quickly process an abundance of complex information and make critical decisions. Research also suggests that the introduction of more narrow or tangible assessment criteria, and the implementation of evaluation processes that allow experts to focus on the information that they value most, might also help counter a reliance on visual information.

Even if "winners" selected under the video-only condition are indeed more successful in the long run, we cannot assume that the primacy of visual information leads to wise decisions. Qualities inferred from visual cues led the original experts to promote and award top prizes to certain people. However, these same visual attributes are also likely to be salient to other audiences and judges. Successive rounds of evaluation may then contribute further to the long-term success of such individuals and groups. It may be possible to reduce the use of visual information at the stage of professional selection yet difficult to eliminate such dependence at the stage of general consumption.

My findings also suggest that the vision heuristic leads people to focus on particular aspects of group performance, namely leadership and group dynamics. In the domain of music, the original expert judges were familiar with the standardized physical arrangement of group members, which indicate clearly which musician is responsible for leading the group. Having identified group leaders, it may have been simple for judges to focus on those individuals. Judges may have assumed that the quality of those individuals was a good indicator of the overall quality of the groups in which they played, given that leadership behavior affects many aspects of group performance (Tschan et al., 2006). For teams in other domains, it may not be as apparent how leadership impacts group performance. Future research could examine whether judgment is more influenced by formalized leadership roles or by the greater salience of visible cues displayed by effective leaders. One possibility is to use a 2×2 design to test the relative effects of role and leadership cues on performance and the judgment of performance.

It would also be valuable for future work to investigate how the most effective leaders influence their teams through nonverbal means to best coordinate performance. Whereas the sight of non-leaders did not allow participants to identify the competition outcomes, the sight of leaders did allow for better identification rates; there may not be a full translation of nonverbal cues (including physical mimicry) from leaders to non-leaders. If the outcome of interest is dependent on the performance of all group members, we may need to more carefully examine whether the lack of professional attention to non-leaders may lead to decision-making costs.

In a world where orchestras spend inordinate amounts of time and money wooing star conductors at the same time that orchestral musicians remain nameless and faceless, the Orpheus Chamber Orchestra has managed to flourish without an obvious leader. This ensemble demonstrates the value of visual communication in group performance. The irony is that although visual communication is used as a *means* of crafting the definitive ensemble sound, musicians actually make professional judgments about ensemble performance through those means alone—not through the sound of the music.

Acknowledgments

The author thanks Teresa Amabile, Sigal Barsade, Max Bazerman, Boris Groysberg, and Jeff Polzer for their insightful feedback and support in the development of the work; and Jim Berry, Katie Shonk, and Bart Vanneste for their methods expertise and wise verse. This research was supported by Harvard University and the Wyss Foundation.

References

Ackerman, J. M., Nocera, C. C., & Bargh, J. A. (2010). Incidental haptic sensations influence social judgments and decisions. *Science, 328*, 1712.

Allport, G. W. (1954). *The nature of prejudice*. New York: Addison Wesley.

Amabile, T., & Kabat, L. (1982). When self-descriptions contradict behavior: Actions do speak louder than words. *Social Cognition, 1*, 311–335.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.). *Advances in experimental social psychology* (Vol. 32, pp. 201–272). San Diego, CA: Academic Press.

Ambady, N., Conner, B., & Hallahan, M. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality & Social Psychology, 77*(3), 538–547.

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*(3), 431–441.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.). *Advances in social cognition* (Vol. 10, pp. 1–48). Mahwah, NJ: Erlbaum.

Bargh, J. A. (2006). What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behaviour. *European Journal of Social Psychology, 36*(2), 147–168.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399). New York: Guilford.

Barsade, S. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47*, 644–675.

Brandimonte, M. A., Schooler, J. W., & Gabbino, P. (1997). Attenuating verbal overshadowing through visual retrieval cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 915–931.

Bryce, W., & Yalch, R. (1993). Hearing versus seeing: A comparison of learning of spoken and pictorial information in television advertising. *Journal of Current Issues and Research in Advertising, 15*(1), 1–20.

Burson, K., Larrick, R., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90*, 60–77.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of person as social entities. *Behavioural Science, 3*, 14–25.

Carney, D. R., Cuddy, A. J. C., & Yap, A. J. (2010). Power poses: Brief nonverbal displays cause neuroendocrine change and increase risk tolerance. *Psychological Science, 21*, 1363–1368.

Carrere, S., & Gottman, J. M. (1999). Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion. *Family Process, 38*, 293–301.

Castellan, N. (1973). Multiple-cue probability learning with irrelevant cues. *Organizational Behavior and Human Performance, 9*(1), 16–29.

Chinander, K., & Schweitzer, M. (2003). The input bias: The misuse of input information in judgments of outcomes. *Organizational Behavior and Human Decision Processes, 91*(2), 243–253.

Curhan, J. R., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology, 92*(3), 802–811.

DePaulo, B. M., Lassiter, G., & Stone, J. I. (1982). Attentional determinants of success at detecting deception and truth. *Personality and Social Psychology Bulletin, 8*(2), 273–279.

Dovidio, J. F., Evans, N. E., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22*, 22–37.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34*, 590–598.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46*, 23–30.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Gilbert, D., & Hixon, J. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509–517.

Gloor, P., Oster, D., & Fischbach, L. (2011). Jazzflow – Analyzing "group flow" among jazz musicians through "honest signals". In *Proceedings of the 3rd international conference*. Basel, Switzerland: Collaborative Innovation Networks COINS.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review, 90*, 715–741.

Greenwald, A. G., McGhee, D. E., & Schwartz, L. K. (1998). Measuring individual differences in implicit cognition: The implicit association task. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Hackman, J. R. (2002). *Leading teams: Setting the stage for great performances*. Boston: Harvard Business School Press.

Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). Englewood Cliffs, NJ: Prentice-Hall.

Hackman, J. R. (2005). Rethinking team leadership *or* Team leaders are not music directors. In D. M. Messick & R. M. Kramer (Eds.), *New directions in the psychology of leadership* (pp. 115–142). Mahwah, NJ: Erlbaum.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review, 103*, 336–355.

Hamilton, D. L., Sherman, S. J., & Maddox, K. B. (1999). Dualities and continua: Implications for understanding perceptions of persons and groups. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 606–626). New York: Guilford Press.

Harrigan, J. A., Wilson, K., & Rosenthal, R. (2004). Detecting state and trait anxiety from auditory and visual cues: A meta-analysis. *Personality and Social Psychology Bulletin, 30*(1), 56–66.

Hoegg, J., & Alba, J. W. (2007). Taste perception: More than meets the tongue. *Journal of Consumer Research, 33*(4), 490–498.

Hsee, C. K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology, 86*(5), 680–695.

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology, 8*(6), R190.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Khodyakov, D. (2007). The complexity of trust-control relationships in creative organizations: Insights from a qualitative analysis of a conductorless orchestra. *Social Forces, 86*(1), 1–22.

Kreuger, L. E. (1973). Effect of irrelevant surrounding material on speed of same-different judgment of two adjacent letters. *Journal of Experimental Psychology, 98*(2), 252–259.

Lamb, G. (2001). Wait, wait, where's the conductor? *Christian Science Monitor*, 13.

Lane, S. M., & Schooler, J. W. (2004). Skimming the surface: Verbal overshadowing of analogical retrieval. *Psychological Science, 15*, 715–719.

Leyens, J., Yzerbyt, V., & Schadron, G. (1992). The social judgeability approach to stereotypes. *European Review of Social Psychology, 3*(1), 91–120.

Malhotra, D., & Bazerman, M. H. (2008). Psychological influence in negotiation: An introduction long overdue. *Journal of Management, 34*(3), 509–531.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.

Melcher, J., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste. *The Journal of Memory and Language, 35*, 231–245.

Mintzberg, H. (1975). The manager's job. *Harvard Business Review, 53*, 49–61.

Murnighan, K., & Conlon, D. (1991). The dynamics of intense work groups: A study of British string quartets. *Administrative Science Quarterly, 36*(2), 165–186.

Nisbett, R. E., & Bellows, N. (1977). Verbal reports about causal influence on social judgment: Private access versus public thesis theories. *Journal of Personality and Social Psychology, 35*, 613–624.

Nisbett, R. E., & Valins, S. (1971). Perceiving the causes of one's own behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Wiener (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 63–78). Morristown, NJ: General Learning Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.

Nisbett, R., Zukier, H., & Lemley, R. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13*(2), 248–277.

Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology, 46*(2), 315–324.

Paivio, A. (1978). A dual coding approach to perception and cognition. In H. L. Pick & E. Salzman (Eds.), *Modes of perceiving and processing information* (pp. 39–51). Hillsdale, NJ: Erlbaum.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to persuasion*. New York: Springer-Verlag.

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science, 19*, 109–111.

Rule, N. O., & Ambady, N. (2011). Face and fortune: Inferences of personality from Managing Partners' faces predict their firms' financial success. *The Leadership Quarterly, 22*, 690–696.

Sager, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology, 39*, 590–598.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 17*, 36–71.

Seifter, H., & Economy, P. (2001). *Leadership ensemble*. New York: Henry Holt.

Sherman, S. J., Castelli, L., & Hamilton, D. L. (2002). The spontaneous use of a group typology as an organizing principle in memory. *Journal of Personality and Social Psychology, 82*, 328–342.

Simon, J., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*(3), 300–304.

Simon, J., & Small, A. (1969). Processing auditory information: Interference from an irrelevant cue. *Journal of Applied Psychology, 53*(5), 433–435.

Smith, E. R. (1984). Model of social inference processes. *Psychological Review, 91*, 392–413.

Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., & de Gelder, B. (2010). I feel your voice: Cultural differences in the multisensory perception of emotion. *Psychological Science, 21*(9), 1259–1262.

Thompson, J. (1967). *Organizations in action*. New York: McGraw-Hill.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623–1626.

Traub, J. (1996). Passing the baton: What CEOs could learn from the Orpheus Chamber Orchestra. *New Yorker, 70*(25), 100–105 (August 26).

Tsay, C. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences, 110*(36), 14580–14585.

Tschan, F., Semmer, N. K., Gautschi, D., Hunziker, P., Spychiger, M., & Marsch, S. (2006). Leading to recovery: Group performance and coordinative activities in medical emergency driven groups. *Human Performance, 19*(3), 277–304.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Valdesolo, P., Ouyang, J., & DeSteno, D. A. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*, 693–695.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181–192.

Wittenbaum, G. M., Vaughan, S. I., & Stasser, G. (1998). Coordination in task-performing groups. In R. S. Tindale et al. (Eds.), *Social psychological applications to social issues. Applications of theory and research on groups* (Vol. 4, pp. 177–204). New York: Plenum Press.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10*, 109–120.

Yzerbyt, V. Y., Schadron, G., Leyens, J., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology, 66*, 48–55.

Zampini, M., & Spence, C. (2004). The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies, 19*(5), 347–363.