DECONSTRUCTING A MUSICAL ILLUSION: POINT-LIGHT REPRESENTATIONS CAPTURE SALIENT PROPERTIES OF IMPACT MOTIONS

Michael Schutz and Michael Kubovy
Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904, USA;
E-mail: schutz@virginia.edu or kubovy@virginia.edu.

ABSTRACT

Although visual information affects auditory perception in a variety of tasks, audition is generally believed to be relatively immune from visual influence when judging tone duration. However, Schutz and Lipscomb (2007) report a musical illusion in which physical gestures influence the perceived duration of notes performed on the marimba. In order to better understand which aspects of these gestures are responsible for the illusion, we created a "schematic marimbist" consisting of either a four-point skeleton or a single moving dot. This schematic abstraction captured the essential properties of the gestures, replicating the effect under both conditions. Therefore, this illusion requires seeing only a sudden change in gesture direction — independent of the depiction of a struck object. As this finding means that it can be replicated with a minimum of visual information, it will be useful in facilitating future research aimed at uncovering the reason for this break with the widely accepted theory of 'optimal integration'.

RESUME

Malgré les effets de l'information visuelle sur la perception auditive lors de plusieurs tâches variées, l'audition est généralement considérée relativement immune à l'influence visuelle lors d'un jugement de la durée d'un son musical. Cependant, Schutz et Lipscomb (2007) décrivent une illusion musicale dans laquelle les gestes physiques influencent la durée perçue de notes jouées sur un marimba. Pour mieux comprendre quelles caractéristiques de ces gestes peuvent être attribuées à cette illusion, nous avons créé un "marimbiste schématique" qui consiste soit d'un bras squelette à quatre points soit d'un seul point animé. Cette abstraction schématique a capté les caractéristiques essentielles des gestes puisque les effets visuels se sont reproduits lors des deux conditions schématiques. La seule condition requise de cette illusion est donc de voir un changement subit dans la direction d'un geste – indépendamment d'une représentation de l'objet frappé. Puisque ces résultats indiquent que l'illusion peut être reproduite avec un minimum d'information visuelle, ils pourront faciliter l'éventuelle recherche visée à la découverte de la cause de cette déviation de la théorie d' "intégration optimale" couramment acceptée.

1. INTRODUCTION

1.0 Music: A multi-modal experience

Since the advent of recording technology, it has been tempting to think of music as a purely auditory phenomenon. It is now apparent, however, that non-auditory information such as kinesthetic feedback (Phillips-Silver & Trainor, 2005, 2007) and visual information (Schutz, 2008; Thompson, Graham, and Russo, 2005) play important roles. In fact, certain aspects of a musical performance can be communicated through purely non-auditory means.

For example, the relative sizes of sung musical intervals can be discerned from a singer's lip movements (Thompson & Russo, 2007). Likewise, a performer's emotional intentions can be inferred by watching their body movements, even in the absence of facial information (Dahl & Friberg, 2007). As both of these studies used vision alone, they effectively demonstrate that it *can* be a salient channel for communicating musical information. However, when determining the degree to which vision plays a meaningful role in the musical experience, it is important to examine whether it alters the listeners' experience. To this end, Thompson, Russo, and Quinto (2008) provide a convincing demonstration. They combined auditory and visual presentations of "happy" (major third) and "sad" (minor third) intervals, and asked participants to judge the emotional tenor (the affect) of each audio-visual pairing. Even though (1) participants were asked to judge auditory information alone and (2) they had to concurrently perform a distracter task designed to minimize "cross-talk" between the modalities, the type of visual information used (happy vs. sad) changed affect ratings.

Vision can affect virtually all aspects of the musical experience: evaluations of expressivity (Davidson, 1993, 1994), audience interest (Broughton & Stevens – in press), judgments of musical tension and phrasing (Vines, Krumhansl, Wanderley, & Levitin, 2006), and even assessments of performance quality (Wapnick, Darrow, Mazza, & Dalrymple, 1997, Wapnick, Mazza, & Darrow, 1998). It can also affect the perception of low-level attributes such as pitch (Thompson et al., 2005, Gillespie, 1997), loudness (Rosenblum & Fowler, 1991), note duration (Schutz & Lipscomb, 2007), and timbre (Saldaña & Rosenblum, 1993). Furthermore, it can improve lyric comprehension (Hidalgo-Barnes & Massaro, 2007), as well as affect judgments of musical dissonance (Thompson et al., 2005, experiment 1), and interval size (Thompson et al., 2005, experiment 3)

Although many musical instruments have been used to study visual influences, percussion offers a particularly rich domain for exploring such issues. Likely, this reflects the relatively large physical motions used by percussionists, and the clear causal relationship between their gestures and sounds (Schutz, 2008). In particular, the marimba (a tuned, wooden bar-percussion instrument similar to the xylophone) has received a great deal of research attention in recent years, including studies showing the importance of visual information with respect to the communication of emotional intention (Dahl & Friberg, 2007, Broughton & Stevens – in press), audience interest (Broughton & Stevens – in press) and note duration (Schutz & Lipscomb, 2007).

1.1 The Schutz-Lipscomb illusion

Schutz and Lipscomb (2007) report an audio-visual illusion in which an expert musician's gestures affect the perceived duration of a note without changing its acoustic length. To demonstrate this, they recorded world-renowned marimbist Michael Burritt (Professor of Percussion at the Eastman School of Music) performing single notes on the marimba using long and short gestures. They paired both types of sounds with both types of gestures, resulting in a combination of natural (i.e., congruent gesture–note pairs) and hybrid (i.e., incongruent gesture–note pairs) stimuli. They informed participants that some auditory and visual components had been mismatched, and asked them to judge tone duration based on the auditory component alone. Despite these instructions, the participants' duration ratings were strongly influenced by visual gesture information (i.e., notes were rated as longer when paired with long gestures than when paired with short gestures). This suggests that the integration of visible striking gestures with heard percussive sounds is perceptually obligatory.

1.2 Why this is puzzling: Previous work on audio-visual integration

These results contradict the view that judgments of tone duration are relatively immune to visual influence (Walker & Scott, 1981, Welch & Warren, 1980), i.e., in temporal tasks visual influence on audition is negligible. For example, audition affects judgments of light duration, but vision does not influence judgments of tone duration (Walker & Scott, 1981). Likewise, the rate of auditory flutter (i.e. number of tones per second) affects the perceived rate of concurrent visual flicker, but the rate of visible flicker either fails to affect the perceived rate of concurrent auditory flutter (Shipley, 1964) or affects it minimally (Welch, DuttonHurt, & Warren, 1986).

Generally, visual information dominates conflicting auditory information only when it is of higher quality — such as in source localization. In the ventriloquism illusion, for example, speech is heard to originate from the moving lips of a silent puppet (Jack & Thurlow, 1973), because the spatial resolution of the visual system is significantly better than that of the auditory system. This dominance is not limited to speech, as shown by similar effects involving non-speech sounds (Bertelson & Radeau, 1981, Bertelson, Vroomen, de Gelder, & Driver, 2000, Jackson, 1953, Thomas, 1941).

This pattern of results led to the formulation of the 'optimal integration hypothesis,' according to which intermodal conflicts are resolved by giving more weight to the modality providing the more reliable information (Ernst & Banks, 2002, Alais & Burr, 2004). This theory has been tested with many different types of cross-modal integration tasks, including visual-haptic (Gepshtein, Burge, Ernst, & Banks, 2005, Guest & Spence, 2003, Miller, 1972) and audio-visual (Alais & Burr, 2004, Ernst & Banks, 2002). It correctly predicts reversal of modality dominance when lowering the quality of information in the generally dominant modality. For example, although the rate of visible flicker does not generally influence the perceived rate of auditory flutter (Shipley, 1964, Welch et al., 1986), vision does exert such an influence when the quality of auditory information is low (Wada, Kitagawa, & Noguchi, 2003).

1.3 Causal relationships

Our previous research has shown, however, that the Schutz-Lipscomb illusion is not based on information quality, but rather on perceived causality. For example, the marimbist's gestures do not affect the perceived duration of non-percussive sounds (such as those produced by a french horn or a clarinet), but they do affect judgments of piano tones (which are also produced by an impact event — that of a hammer striking a string). Likewise, when the causal relationship between the auditory and the visual streams is temporally disrupted (e.g., the percussive sound precedes the visible impact) the gestures fail to influence auditory perception. Furthermore, this manipulation is asynchronous – sounds lagging the moment of impact continue to be influenced by the gestures despite the lack of influence when leading by equal amounts. This is consistent with the physical structure of our environment, in which the speed of sound is substantially less than that of light (Schutz & Kubovy, in press).

These results suggest a causal account of the illusion – gestures integrate (and therefore influence) only the sounds they could have caused. Additionally, they argue against a post-perceptual account (alternatively referred to as a response-bias or cognitive correction), by showing that the illusion is not an artifact of differential amounts of motion between the long and short gestures. Furthermore, they provide a clear way to test whether the illusion is in fact consistent with the optimal integration hypothesis, which in this paradigm predicts that visual influence is related to auditory ambiguity.

In the literature, stimulus ambiguity is operationally measured by response variability. To this end, we compared the variability of duration ratings for percussive sounds to those of the non-percussive sustained sounds (such as the french horn and clarinet) from the previous experiment. The results were inconsistent with the optimal integration hypothesis — ratings were no more variable for the visually influenced (percussive) than the non-influenced (sustained) tones. Furthermore, the visual influence on percussive tones was not related to decay time: slowly decaying sounds were no more influenced than quickly decaying ones (Schutz & Kubovy, in press).

Our interest in the role of causality in cross modal integration is not without precedent. Previous research on issues related to the "unity assumption" (Welch, 1972; see also: Spence, 2007, Welch and Warren, 1980, Welch, 1999, Vatakis and Spence, 2008, Vroomen, 1999) and the "identity decision" (Bedford, 2001a, 2001b, 2004) represent similar thinking. Such work explores the conditions under which cross-modal influences can occur, and suggests that integration requires a mechanism for inferring that auditory and visual information originate from the same event. We agree that such an "identity decision" is a requirement for integration, and posit that causality serves as one of the primary cues by which it can be triggered.

1.4 Motivation for this study

It is well known that point-light displays are capable of conveying biological motion (Johansson, 1973), and can be effectively used in studies of cross-modal interactions (Arrighi, Alais, & Burr, 2006, Saygin, Driver, & de Sa, 2008). Therefore, we were curious whether they could be used within this paradigm as well. Our motivations were two-fold: (1) to better understand the particular visual cues driving this illusion, and (2) to facilitate future research by establishing the feasibility of using abstractions offering obvious methodological advantages with respect to manipulability and control.

2. METHODS AND PROCEDURE

The original videos showed percussionist Michael Burritt performing notes using long and short gestures on the marimba. They were made with a Cannon GL1 video camera and Audio-Technica AT4041 microphones. Here, we used two types of gestures (long and short) taken from the original stimuli, presenting them at one of three levels of abstraction (shown in Figure 1) for a total of six visual stimuli. The conditions were:

1. Video. Movies displaying the marimbist performing single notes at 3 pitch levels using either long or short gestures.
2. Skeleton. A point-light skeleton version of these videos consisting of 4 white dots connected by line-segments against a black background. The 4 dots tracked the shoulder, elbow, hand, and mallet head.
3. Dot. A reduced version of the point-light skeleton showing one dot, tracking the mallet head.

2.1 Stimuli

We made the animations by tracking the horizontal and vertical location of four key points (shoulder, elbow, hand, and mallet tip) from the original videos on a frame-by-frame basis, using the program GraphClick[1]. Each of these joint locations was rendered as a white dot, with consecutive joints connected by white lines (bottom left panel of Figure 1). These animations were generated in real time on a trial-by-trial basis using custom designed software[2]. The animations contained no further information, and therefore the struck object (originally a marimba bar) was not represented. However, from the motion of the striking implement alone it was clear to us (and to participants in a pilot experiment) that the motion represented an impact event.

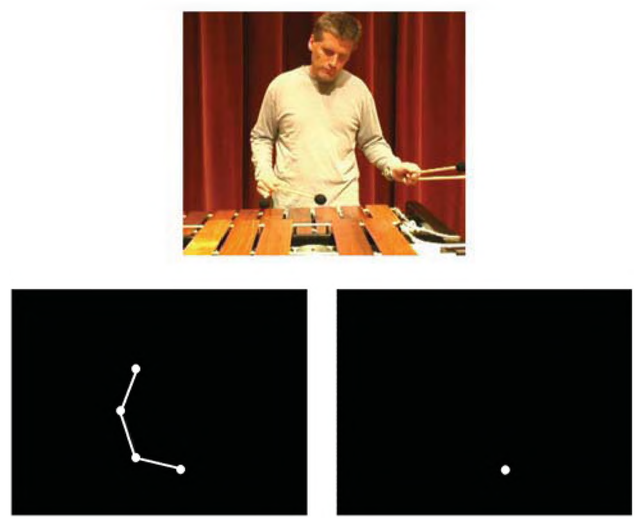

Figure 1:
Top: The video stimuli, consisting of videos displaying a marimbist playing with long and short gestures.
Bottom left: The skeleton stimuli, consisting of 4-point versions of the video stimuli. The points tracked the performer's shoulder, elbow, hand, and mallet head.
Bottom right: The single point dot stimuli, tracking only the position of the mallet head.

For auditory stimuli we used natural and damped marimba tones from the three pitch levels in the original experiment. For the natural tones, the marimbist allowed the sound to decay naturally, whereas for damped tones, he manually damped the bar after striking. One example of each type (damped, natural) was chosen for three pitch levels: E1 (~82 Hz), D4 (~587 Hz), and G5 (~1568 Hz), for a total of six auditory stimuli. We crossed (i.e., took the

[1] http://www.arizona-software.ch/graphclick

[2] Designed and implemented by Simeon Fitch of Mustard Seed Software http://www.mseedsoft.com

Cartesian product of) these six visual and six auditory streams, to produce 36 stimuli (samples of the stimuli used can be viewed online at http://sites.google.com/site/schutzresearch/virtualmarimbist).

2.2 Procedure

Thirty-eight participants[3] (receiving course credit) saw the stimuli on a computer screen, and rated the duration of the tone under two conditions: audio-visual and audio-alone. They rated the duration of each using an on-screen slider with endpoints labeled "Short" and "Long." We translated their rating for each trial into an integer in the [0, 100] interval, where 0 stood for short. In the audio-visual condition, they were instructed to base their judgments on the auditory information alone.

To discourage participants from ignoring the screen, we asked them to respond to a second question concerning the level of agreement between the auditory and visual components, using a second slider with endpoints "Low agreement" and "High agreement." Previous research has shown that asking participants to provide agreement ratings does not interfere with the primary task (Rosenblum & Fowler, 1991, Saldaña & Rosenblum, 1993), a finding that has held throughout our previous work (Schutz & Lipscomb, 2007, Schutz & Kubovy, in press). As the purpose of this agreement rating was only to force attention to both modalities, they were not analyzed and will not be discussed further.

After a warm-up block containing a random selection of stimuli, we presented each of the 36 audio-visual stimuli three times. The 108 trials were organized into three blocks, with each block containing one instance of each stimulus (within-block trial order was randomized independently).

3. RESULTS AND DISCUSSION

We analyzed the duration ratings using mixed-effects linear models in which all manipulated variables were treated as fixed effects within participants. Several textbooks (Baayen, 2008, Kreft & Leeuw, 1998, Raudenbush & Bryk, 2002, Snijders & Bosker, 1999) present mixed-effects analyses, which have considerable advantages (Baayen, Davidson, and Bates, in press and Maxwell & Delaney, 2004, Part IV). We report each result in terms of an effect (and its standard error, SE, in parentheses), from which a Cohen effect size, d, can be obtained by dividing the effect by its SE. To these we added a 95% confidence interval, as well as a p-value for a test of the null hypothesis that the effect in question is 0.

The single-point animations replicated the illusion (Figure 2). Overall, visual information affected the duration ratings by an estimated 9.6 (±2.0) rating points (95% CI: [5.6, 13.6], $p \approx 0$ [4]). We found no meaningful difference between the visual influence from the full videos and the influence from the 4-point animations (2.8±2.1, 95% CI: [-1.3, 6.8], p=0.2). Similarly, we found no meaningful difference between influence from the full videos and the 1-point animations (1.4±2.0, 95% CI: [-2.6, 5.4], p=0.5). As expected, natural notes were judged longer than damped (by 9.2±2.2, 95% CI: [4.9,13.5], p=0.004). However, this parameter had no effect on the degree of visual influence, as shown by the non-significant interaction between note type and visual stroke type (0.45±2.9, 95% CI: [−5.3, 6.1], p=0.9). This is consistent with the results of Schutz and Lipscomb (2007), as well as our subsequent investigations (Schutz & Kubovy, in press).

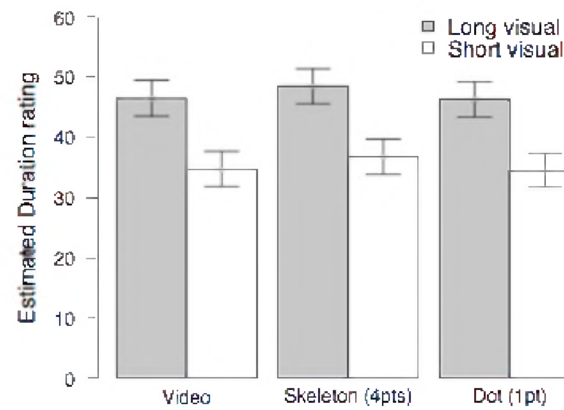

Figure 2: Visual influence was consistent across presentation conditions. The difference in ratings for notes paired with long impact gestures (grey) and short gestures (white) was similar across the video, 4-point skeleton animations, and 1-point dot animation conditions. Error bars indicate a 95% confidence interval.

4. CONCLUSIONS

Given the success of visual (Johansson, 1973) and audio-visual (Saygin et al., 2008) point-light displays in capturing biological motion, we were not surprised to see that the 4-point skeleton produced patterns of visual influence similar to those obtained in the video condition. However, we were pleased to discover that a single dot was sufficient to replicate the original effect (in a different experiment using only the single-dot animations, most participants reported informally that the motion appeared to depict an impact). These results attest to the robustness of the illusion, suggesting interesting possibilities for future research. Furthermore, they are consistent with our conjecture that a crucial component of the illusion is the perception of an audio-visual causal relationship.

We believe these results will be of potential interest for at least two reasons. First, because the illusion has already proven informative with respect to our understanding of

[3] Because previous research within this paradigm found no meaningful difference between musically trained (Schutz & Lipscomb, 2007) and musically untrained (Schutz & Kubovy, in press) participants, we did not record any information regarding years of musical training.

[4] Here we use ≈ 0 to indicate that within the computational precision available, for all practical purposes the value is not meaningfully different from 0.

sensory integration, we believe there will be interest in its further study. Consequently, reducing the visual component of the illusion to a single dot will allow us to carry out controlled investigations to discover which aspects of the visual stimulus are responsible for the illusion. This will prove valuable in understanding how the perceptual system integrates sensory information across modalities, and further explore the role of causality as a cue for triggering the identity decision.

Second, we believe that a greater understanding of and appreciation for the role of visual information in shaping the musical experience holds clear artistic value. Given that some expert musicians use vision to strategically enhance audience interest (Broughton & Stevens – in press), communicate musical structure (Vines et al., 2006), and manipulate evaluations of performance quality (Wapnick et al., 1998), deepening our understanding of this process will be helpful to both performers and audiences alike. Future experiments using abstract versions of the original striking gestures will help illuminate how to best use this information to enhance the quality of performer-audience communication.

ACKNOWLEDGMENTS

We thank S. Fitch of Mustard Seed Software for designing the program used to run the experiment, and J. Armontrout for his help in producing stimuli and running participants. This research was supported in part by a grant from NIDCD R01 DC 005636 (Michael Kubovy, PI).

REFERENCES

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262.

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, *6*, 260–268.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Bedford, F. L. (2001a). Object identity theory and the nature of general laws. *Cahiers de Psychologie Cognitive—Current Psychology of Cognition*, *20*, 277–293.

Bedford, F. L. (2001b). Towards a general law of numerical/object identity. *Cahiers de Psychologie Cognitive—Current Psychology of Cognition*, *20*, 113–175.

Bedford, F. L. (2004). Analysis of a constraint on perception, cognition, and development: One object, one place, one time. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 907–912.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*, 578–584.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*, 321–332.

Broughton, M., & Stevens, C. (in press). Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*(5), 433-454.

Davidson, J. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*.

Davidson, J. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies*, *26*, 279–301.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Gepshtein, S., Burge, J., Ernst, M. O., & Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity. *Journal of Vision*, *5*, 1013–1023.

Gillespie, R. (1997). Ratings of violin and viola vibrato performance in audio-only and audiovisual presentations. *Journal of Research in Music Education*, *45*, 212–220.

Guest, S., & Spence, C. (2003). Tactile dominance in speeded discrimination of textures. *Experimental Brain Research*, *150*, 207–207.

Hidalgo-Barnes, M., & Massaro, D. W. (2007). Read my lips: an animated face helps communicate musical lyrics. *Psychomusicology*, *19*, 3-12.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills*, *37*, 967–979.

Jackson, C. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, *5*, 52–65.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *2*, 201–211.

Kreft, I., & Leeuw, J. de. (1998). *Introducing multilevel modeling*. London, UK: Sage.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ, USA: Erlbaum.

Miller, E. A. (1972). Interaction of vision and touch in conflict and nonconflict form perception tasks. *Journal of Experiment Psychology*, *96*, 114–123.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*, 1430.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*, 533-546.

Raudenbush, S. W., & Bryk, A. S. (2002). *linear mode: Applications and data analysis methods*. London, UK: Sage.

Rosenblum, L. D., & Fowler, C. A. (1991). Audiovisual investigation of the loudness-effort effect for speech and nonspeech events. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 976–985.

Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics*, *54*, 406–416.

Saygin, A. P., Driver, J., & de Sa, V. R. (2008). In the footsteps of biological motion and multisensory perception. *Psychological Science*, *19*, 469–475.

Schutz, M. (2008). Seeing music? What musicians need to know about vision. *Empirical Musicology Review*, *3*, 83–108.

Schutz, M., & Kubovy, M. (in press). Causality and cross-modal integration. *Journal of Experimental Psychology: Human Perception and Performance*.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: vision influences perceived tone duration. *Perception*, *36*, 888–897.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science*, *145*, 1328–1330.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, UK: Sage.

Spence, C. (2007). Audiovisual multisensory integration. *Acoustic Science & Technology*, *28*, 61–71.

Thomas, G. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology*, *28*, 163–175.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *156*, 203–227.

Thompson, W. F., & Russo, F. A. (2007). Facing the music. *Psychological Science*, *18*, 756–757.

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition & Emotion*, *22*, 1457–1470.

Vatakis, A., & Spence, C. (2008). Evaluating the influence of the 'unity assumption' on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica*, *127*, 12–23.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross–modal interactions in the perception of musical performance. *Cognition*, *101*, 80–113.

Vroomen, J. (1999). Ventriloquism and the nature of the unity decision. Commentary on Welch. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 389–393). Amsterdam: Elsevier.

Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio-visual integration in temporal perception. *Psychophysiology*, *50*, 117–124.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1327–1339.

Wapnick, J., Darrow, A.-A., Mazza, J. K., & Dalrymple, L. (1997). Effects of physical attractiveness on evaluation of vocal performance. *Journal of Research in Music Education*, *45*, 4270-479.

Wapnick, J., Mazza, J. K., & Darrow, A.-A. (1998). Effects of performer attractiveness, stage behavior, and dress on violin performance evaluation. *Journal of Research in Music Education*, *46*, 510–521.

Welch, R. B. (1972). The effect of experienced limb identity upon adaptation to simulated displacement of the visual field. *Perception & Psychophysics*, *12*, 453–456.

Welch, R. B. (1999). Meaning, attention, and the 'unity assumption' in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 371–387). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.

Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics*, *39*, 294–300.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638–667.