HOW DO PEOPLE *THINK* THEY REMEMBER MELODIES AND TIMBRES? PHENOMENOLOGICAL REPORTS OF MEMORY FOR NONVERBAL SOUNDS

Michael A. Nees, Joanna Harris, & Peri Leong

Lafayette College,
Department of Psychology,
Oechsle Hall, Easton, PA 18042, USA
[neesm][harrisje][leongp]@lafayette.edu

ABSTRACT

Memory for nonverbal sounds such as those used in sonifications has been recognized as a priority for cognitive-perceptual research in the field of auditory display. Yet memory processes for nonverbal sounds are not well understood, and existing theory and research have not provided a consensus on a mechanism of memory for nonverbal sounds. We report a new analysis of a qualitative question that asked participants to report the strategy they used to retain nonverbal sounds—both melodies and sounds discriminable primarily by timbre. The question was originally posed as part of the debriefing procedure for three separate memory experiments whose primary findings are reported elsewhere. Results of this new analysis suggested that auditory memory strategies—remembering acoustic properties of sounds—were common across both types of sounds but were more commonly reported for remembering melodies. Motor strategies were also more frequently reported for remembering melodies. Both verbal labeling of sounds and associative strategies—linking the sounds to existing information in memory—were more commonly reported as strategies for remembering sounds discriminable primarily by timbre. Implications for theory and future research are discussed.

1. INTRODUCTION

Kramer [1] identified the "absence of persistence" as a potential weakness of auditory displays. Unlike their visual counterparts, auditory displays are transient—a quality that may increase demands on memory, particularly if comparisons of sounds over time are required to accomplish tasks with auditory displays. Accordingly, the seminal Sonification Report [2] identified studies of memory for sounds as a priority in a research agenda for perceptual and cognitive scientists working in the field of auditory display. More specifically, Flowers [3] noted that sonifications should be designed to accommodate the limitations of auditory memory processes, including sensory memory and working memory.

Nearly two decades later, important research questions remain unresolved regarding memory for sounds—especially nonverbal sounds such as those used in sonifications. Influential theories of attentional and memory processes (e.g., [4], [5]) have little to say about cognitive processes for nonverbal sounds. This omission may result, at least in part, from the lack of certainty about basic representations of nonspeech sounds in memory.

1.1. Auditory Sensory Memory and Working Memory

There is relatively wide agreement that sounds persist in memory for a brief period following stimulation. This phenomenon has been called *auditory sensory memory* or *echoic memory* in the literature. Auditory sensory memory preserves a high-fidelity mental representation of sounds in an auditory format—that is, for the duration of auditory sensory memory, sounds are accessible in memory as sounds per se. Further, auditory sensory memory processes do not seem to require attentional resources. The precise duration of auditory sensory memory is not currently known. Estimates have ranged widely—from a second or two in some studies up to 30 seconds or longer in others (for a review, see [6]).

Unlike the brief, passive auditory sensory store, *working memory* entails the active maintenance, rehearsal, and processing of information, and this active processing can sustain information in working memory indefinitely (see [6]–[8]). Verbal working memory, including auditory language, is widely believed to involve a subvocal (i.e., silent or covert) articulatory mechanism [4]. A miscellany of theoretical perspectives has emerged on the working memory rehearsal mechanism for nonverbal sounds. Baddeley and Logie suggested that [9] the working memory mechanism for verbal and nonverbal sounds might be one and the same. Berz [10] proposed a musical working memory mechanism that functioned with independence from verbal working memory processes. Attention-based rehearsal has been speculated to play a role in memory for timbre [11]—a property of sounds that some have purported is impossible to rehearse via subvocal articulation [12]. Yet other researchers have endorsed the perspective that nonverbal sounds cannot be rehearsed and are instead remembered "automatically" [13], although evidence to the contrary has been presented (e.g., [14]).

1.2. Encoding Strategies

Divergent hypotheses regarding the mechanism of working memory for nonverbal sounds may persist in part due to variability in the encoding strategies people use to remember sounds. People exhibit individual differences in the encoding strategies they use to remember perceptual stimuli, and these differences can be observed in behavioral studies of task performance. Macleod, Hunt, and Mathews [15], for example, exposed two distinct strategies used by participants in sentence-picture verification tasks. These tasks require participants to encode a sentence or phrase in memory (e.g., "plus is above star."). The time required to encode the stimulus is recorded as a dependent variable—comprehension

time. Following encoding, participants view a picture that either matches (e.g., a plus above a star) or does not match (e.g., a star above a plus) the state described in the sentence. Participants make a speeded response to indicate whether the sentence and picture matched, and the time required to respond is recorded as another dependent variable—verification time. Macleod et al. examined individual patterns of comprehension times and verification times and identified two different strategies for encoding the sentence. Participants who used a verbal encoding strategy remembered the sentence as words. Their characteristic pattern of task performance was short comprehension times and relatively longer verification times—presumably because they had to transform the sentence from its verbal memory code into a pictorial code for comparison during verification. Participants who used a visual imagery strategy exhibited long comprehension times—presumably because they were transforming the sentence to a picture during the comprehension phase of trials. Their verifications time were faster, however, as they could quickly compare their pictorial mental image to the picture stimulus. The individual difference that determined the strategy adopted seemed to be spatial ability; participants with higher scores on a psychometric test of spatial ability tended to adopt a visual imagery strategy. A later study [16] showed that participants could be trained to successfully adopt either strategy, regardless of their preferred default strategy.

1.3. Multiple Encoding Strategies for Melodies

Similarly, research on nonverbal sounds—melodies in particular—has suggested that people use multiple encoding strategies to remember tones. Zatorre and Beckett [17] reported that people with absolute pitch encode notes using note names (a verbal encoding strategy), motor codes (e.g., how to produce the note on a musical instrument), auditory codes (i.e., remembering sounds as sounds per se), and visual imagery (e.g., remembering the note by creating a mental image of its representation on a musical staff). Mikumo's (e.g., [18]) research suggested that these four codes may be widely available as encoding strategies, even for listeners who do not possess absolute pitch.

Nees and colleagues [19], [20] extended the sentence-picture verification task to include sound stimuli—brief, two-note nonspeech sounds that were like sonifications. Nees and Walker [19] demonstrated that tonal stimuli could be encoded as sounds (i.e., an auditory imagery strategy), as visuospatial images (i.e., a visual imagery strategy), or as words (a verbal strategy) depending upon the study instructions (also see [21]). The same study suggested that—for sounds only—the auditory sensory memory trace lingered for at least 3 seconds even when the stimulus had been recoded (e.g., to a verbal or visuospatial representation) in working memory.

1.4. Memory for Timbre

Although a good deal of research has suggested that a variety of encoding strategies can be used to remember melodies and pitched stimuli, the evidence regarding encoding strategies for sounds that are discriminable primarily by timbre is relatively scant. Crowder [22] provided evidence that timbre is stored in memory and can be imagined in the absence of a stimulus, presumably by reinstating representations of timbre held in long term memory. Crowder suggested that, to explain encoding of timbre, "an appeal to sensory rather than motor imagery is justified" (pp. 478) on the grounds that people cannot physically reproduce timbral differences in sounds.

Our inability to physically reproduce sounds was assumed but not tested by Crowder. If correct, this would represent a difference between pitch and timbre memory, because pitch information can be rehearsed using a humming or singing encoding strategy.

Golubock and Janata [23] concluded that memory capacity for timbre was lower than verbal working memory capacity, as verbal items allowed access to both articulatory rehearsal and long-term associations. The specific auditory working memory capacity of sounds discriminable primarily by timbre was found to be just one to two items, in comparison to the typical verbal working memory limit of three to five items. Studies also have shown that the working memory capacity of sounds varying on timbre increases with more diverse stimuli [11], [23]—a finding which suggested that confusability among items with similar timbres could be a source of difficulty in memory for timbre.

1.5. The Current Study

Despite awareness in the auditory display community of the importance of memory for sounds, many questions remain regarding how people remember sounds—especially nonverbal sounds like those used in sonifications. The current study examined subjective reports of strategies used to encode melodies and timbre. Nees and Walker [24] previously reported a study on subjective reports of strategy use during data analysis tasks (such as point estimation) with auditory graphs—a type of sonification that uses pitch changes in time to convey information (for a review, see [25]). They found that, although verbal, visuospatial, auditory, and motor encoding strategies were reported, of the four, verbal strategies were reported most frequently. Further, each of these strategies was reported less frequently than other strategies (e.g., counting in time to the auditory graph and using the auditory graph's contextual auditory cues) that were specific to the data analysis task (rather than representative of a general encoding strategy for sounds). The general approach used by Nees and Walker was adopted here, although the tasks of the current analyses reflect demands that were more purely based in memory than the tasks used by Nees and Walker.

The data analyzed here were collected as part of a debriefing procedure across three previous experiments whose primary results are reported elsewhere (see below). Participants gave an open-ended response indicating the strategy they had used to retain either melodies or the timbre of sounds during an 8 second retention interval. This exploratory investigation sought to examine participants' subjective impressions of how they rehearsed and retained nonverbal sounds in memory during these experiments, and also to examine potential differences in these impressions when the target stimuli were pitched melodies as compared to sounds discriminable by timbre only.

2. METHOD

The data reported here were collected as part of a qualitative post-experimental follow-up question in three separate experiments. The primary results of these experiments are reported elsewhere. The first two experiments (reported in [26]) examined memory for melodies, and the third experiment [27] examined memory for timbre. Although the focus of the current report is the analysis of a single qualitative follow-up question (see section 2.3), for context a brief description of the methods of each experiment follows.

2.1. Memory for Melodies Experiments

Nees, Corrini, Leong, and Harris [26] conducted a pair of studies that were designed to examine both articulatory processes and attentional processes as potential mechanisms of rehearsal of melodies in working memory. The stimuli and task were modeled after the task used by Schendel and Palmer [28]. Participants heard brief four-note melodies, with each note randomly selected from the octave ranging from D4 to E5. Notes were created in the MIDI piano timbre and were 500 ms in duration, including 10 ms onset and offset ramps; each melody was 2000 ms in total duration. On a trial, participants heard a standard melody. Following an 8 second delay, participants heard a second probe melody and indicated whether the probe melody was the same as the initial standard melody. Half of the trials in the study featured the same melody, and half featured different melodies. Different melodies were altered such that either the second or third note in the melody was either raised or lowered by one note. Across four blocks, participants experienced a set of distractor conditions intended to suppress articulatory rehearsal (reading solved math problems aloud), attentional rehearsal (solving math problems and responding manually instead of vocally), or both (reading and solving math problems aloud). In the first experiment, articulatory rehearsal suppression was spoken aloud, whereas in the second experiment, it was silently voiced. Results of both experiments showed that articulatory rehearsal suppression interfered with memory for melodies, whereas interference with attentional rehearsal did not. The studies were interpreted to indicate that articulatory rehearsal plays a role in working memory for melodies.

2.2. Memory for Timbre Experiment

Nees, Leong, and Harris [27] conducted an experiment to follow-up the melodies experiment described above. The stimuli for this experiment were 8 sounds culled from the stimulus set reported in Golubock and Janata ([23], Experiment 1, pp. 402). These synthesized stimuli were designed to be abstract and discriminable by timbre. Golubock and Janata began with a sound that had a fundamental frequency of 341 Hz and 19 harmonics. Next, they systematically altered the attack, spectral centroid, and spectral flux of the sound to create a set of stimuli that were placed in a three-dimensional (attack, centroid, flux) stimulus space. Finally, they conducted listening tests and confirmed that the sounds at the corners of their design space were discriminable. We used the 8 corner stimuli from their stimulus space in our study. Although the sounds were discriminable, they were designed such that discrimination would not be possible using cues related to pitch or existing associations with instruments or real-world sound sources.

The Nees et al. timbre experiment [27] followed the same basic procedure as the melodies experiments described above, except that participants heard only one abstract sound and one probe. The primary timbre experiment found no difference across the interference conditions. Participants exhibited much higher d' (sensitivity) scores in this experiment as compared to the melodies experiments above. We concluded that the task exhibited a form of ceiling effect, and follow-up studies are underway with longer sequences of abstract sounds.

2.3. The Current Study: Reports of Strategy Use

2.3.1. Participants

There were 40 participants in the first melody study (30 females, M age = 19.58 years, SD = 1.22), 36 participants in the second melody study (28 females, M age = 19.53 years, SD = 0.91), and 44 participants (34 females, M age = 19.41 years, SD = 0.87) in the timbre study. Data from one participant in the first melody study were not usable, so the final analysis reported here involved a total sample of N = 119 cases. In all 3 studies, participants rated their musical ability on a scale from 1 ("I have no musical ability at all") to 7 ("I am a professional musician"). The ratings across the two melodies studies and the timbre study were M = 2.85 (SD = 1.44, mdn = 3.00, mode = 2), M = 2.78 (SD = 1.44, mdn = 3.00, mode = 1), and M = 2.68 (SD = 1.54, mdn = 2.00, mode = 1), respectively, which suggested that participants were overwhelmingly non-musicians.

2.3.2. Strategy Use Query

At the end of each of the studies described in 2.1 and 2.2, participants responded to the following question:

"Please briefly describe the strategy you used to remember the [sounds]. If you used different strategies for different parts of the study, describe the different strategies."

The question was open-ended. Participants typed their answers into a text box on a Qualtrics questionnaire. They could type as much or as little as they wished.

2.4. Coding Rubric

The coding rubric for this analysis was adapted from the coding scheme used in Nees and Walker [24]. Nees and Walker's task and stimuli were different from the tasks and stimuli used here, so some modifications to their coding scheme were required. The current rubric adopted five of the strategies described in their rubric directly (verbal, visuospatial, motor, auditory-musical, and differential strategies). Three of the strategies they described were specific to their tasks and stimuli (context, counting, and arithmetic strategies), and we replaced those categories with strategies that were relevant to our tasks and stimuli (i.e., the associative and distraction avoidance strategy categories, and also the no strategy category). The rubric's operational definitions of statements that fell into each strategy category are described below.

2.4.1. Verbal Strategy

A response was coded as indicating a verbal encoding strategy if the participant reported naming sounds, labeling sounds (e.g., any mention of musical note names), comparing sounds to one another using assigned verbal labels (such as "high" and "low"), or mentioning specific sounds by an assigned name. This strategy included any indication that the subject labeled specific sounds with a verbal (i.e., linguistic) tag.

2.4.2. Visuospatial Strategy

A response was coded as indicating a visuospatial encoding strategy if the participant reported that she mentally drew a

picture, used visual imagery, or created a graph her head. Responses that suggested features characteristic of a visual image (slope, line, top, bottom etc.) were also included.

2.4.3. Motor Strategy

A response was coded as indicating a motor encoding strategy if the participant reported a non-articulatory movement strategy (meaning use of the hands or feet, etc.) to assist memory. This strategy included counting on fingers, tapping, moving physically with the mouse or finger or "drawing" on the desk with a finger to remember contour.

2.4.4. Auditory-musical Strategy

A response was coded as indicating an auditory-musical encoding strategy if the participant reported humming, whistling, vocalizing, or imitating, either overtly or covertly, the melody, pitch, or timbre of any part of the stimulus. This strategy included an indication that the participant was attempting to maintain some isomorphic representation of the sound or replay the sound stimulus. Any mention of the participant "hearing" the sound, recording the sound in her mind, or "focusing" on specific notes, etc., also was included in this category.

2.4.5. Associative Strategy

A response was coded as indicating an associative encoding strategy if the participant reported making connections with any information that was already learned. This strategy included associating sounds with existing sounds or other information in long-term memory and use of mnemonics.

2.4.6. Differential Strategy

A response was coded as indicating a differential encoding strategy if the participant reported differences in strategy use based on stimulus differences or experimental conditions. This strategy included responses that offered any indication that the strategy switched over the course of the experiment.

2.4.7. Distraction Avoidance Strategy

A response was coded as indicating a distraction avoidance encoding strategy if the participant reported trying to ignore the interference tasks or to control attention to minimize distraction and maximize focus on the sounds. This strategy included responses indicating that the participant closed her eyes to avoid visual distractions, etc.

2.4.8. No Strategy or No Knowledge of Strategy

A response was coded as indicating no encoding strategy or no knowledge of the strategy if the participant reported having no knowledge of her strategy or indicated that she used no strategy at all.

2.5. Coding Procedure

The coding procedure involved three steps. First, all authors examined all participant responses for fit with the rubric used by Nees and Walker [24] without performing any formal ratings. Second, in a meeting, all three authors discussed the Nees and Walker rubric and its rating criteria. During the meeting, we arrived at a consensus about the applicability of each strategy to our data and adapted some categories of the rubric as described above. Finally, two raters (the second and third authors) independently coded the presence or absence of each strategy in each response across all cases. Coding of strategies was not mutually exclusive—a given response could be categorized as meeting the criteria for more than one strategy.

3.   RESULTS

3.1. Length of Responses

All responses were coded for length using the word count feature of Microsoft Word. The mean length of responses was $M = 25.68$ words ($SD = 25.37$).

3.2. Inter-rater Agreement

The coding scheme was checked for inter-rater agreement among the two raters using a percent agreement measure. Percent agreement was calculated to account for the number of discrepant ratings (i.e. that showed lack of agreement in coding) across all coded categories for the $N = 119$ cases as shown in Equation 1.

$$Percent\ agreement = \frac{119 - total\ discrepancies}{119} \quad (1)$$

Percent agreement was high across the raters (ranging from 92% to 98% agreement for the strategies); thus, no revision to the coding scheme or re-rating was deemed necessary. The final determination of how to code the small number of discrepant ratings was settled by discussion among the raters to arrive at consensus.

3.3. Overall Use of Strategies Across All Cases

Table 1 shows the overall reported use of each strategy across all cases. Notable results here included the high percent of respondents that used an auditory-musical strategy and the low percent of participants that reported having no strategy or no knowledge of their strategy. Most participants, then, apparently could access and report upon a specific approach they used to remember sounds. Perhaps not surprisingly, most reported using a strategy that involved trying to remember the sounds as sounds per se. Verbatim representative examples of the types of statements that fell into each strategy category are given below.

Strategy	Overall Percent Reported (N = 119)
Verbal	18
Visuospatial	12
Motor	16
Auditory-musical	61
Associative	13
Differential	8
Distraction Avoidance	12
No strategy/ knowledge	5

Table 1. Percent of participants reporting using each strategy.

3.3.1. Verbal Strategy

Across all cases, 18 percent of participants reported using a verbal strategy. Representative examples of statements indicating a verbal encoding strategy are shown below.

"*I tried to think of/remember whether the noise was high, low, or in the middle…*"

"*I'd say quick words to the pitch and anticipate the second round with those words.*"

3.3.2. Visuospatial Strategy

Across all cases, 12 percent of participants reported using a visuospatial strategy. Representative examples of statements indicating a visuospatial encoding strategy are shown below.

"*I visualized the notes on a scale as if I was reading music.*"

"*I counted the beats and graphed them in my head in almost a bar graph.*"

3.3.3. Motor Strategy

Across all cases, 16 percent of participants reported using a motor strategy. Representative examples of statements indicating a motor encoding strategy are shown below.

"*I used my fingers to tap it on different locations on the mousepad. The locations corresponded to different pitches.*"

"*Tapping out the music on my fingers.*"

"*I would look up if the note was high and down if the note was low and followed the beat with my eyes.*"

3.3.4. Auditory-musical Strategy

Across all cases, 61 percent of participants reported using an auditory-musical strategy. Representative examples of statements indicating an auditory-musical encoding strategy are shown below.

"*I sung it in my head during the pause.*"

"*I just repeated the melodies in my head as I answered the problems.*"

3.3.5. Associative Strategy

Across all cases, 13 percent of participants reported using an associative strategy. Representative examples of statements indicating an associative encoding strategy are shown below.

"*I associated the noises with different sound effects.*"

"*Sounds reminded me of certain things. One sounded like an instrument being aggressively plucked, another like an angry pop up, another sounded like it got cut off, etc.*"

3.3.6. Differential Strategy

Across all cases, 8 percent of participants reported using a differential strategy. Representative examples of statements indicating a differential strategy are shown below.

"*For the sections with the math problem, I would try to play out the melody like I was playing it on the piano. For the section without the math problems, I tried to visualize the melody on the staff then repeat it 3 times before it played again.*"

"*During the portions with math I tried to read the math as fast as possible so I could keep trying to repeat the melody in my head with little interruption. During the portion without math I tried to repeat the melody and the high low beats as much as I could before the second melody played.*"

3.3.7. Distraction Avoidance Strategy

Across all cases, 12 percent of participants reported using a distraction avoidance strategy. Representative examples of statements indicating a distraction avoidance strategy are shown below.

"*For the last part of the study with the 8 second gap in between each sound, I closed my eyes when listening to the sounds so that I would have no visual distractions.*"

"*I tried to pay as little attention as possible to the math problems.*"

3.3.8. No Strategy or No Knowledge of Strategy

Across all cases, 5 percent of participants reported having no strategy. Representative examples of statements indicating no knowledge of strategy are shown below.

"*I did not really have a strategy.*"

"*I tried to listen carefully but I did not have a strategy.*"

3.4. Comparison of Strategy Use Across Melody and Timbre Stimuli

Chi-square tests (see [29], [30]) compared the proportions of respondents that used each strategy to remember melodies and timbres (see Footnote 1). The results of these analyses are presented in Table 2.

Results showed that participants were significantly more likely to report having used motor and auditory-musical strategies to remember melodies as compared to timbre. For

timbre, participants were significantly more likely to report having used verbal or associative encoding strategies.

Strategy	Melody Percent (N = 75)	Timbre Percent (N = 44)	χ2 (1 df)	p
Verbal	9	32	10.07	.002*
Visuospatial	16	5	3.17	.08
Motor	23	5	6.52	.01*
Auditory-musical	71	45	7.84	.005*
Associative	3	32	19.53	<.001*
Differential	9	7	0.15	.70
Distraction Avoidance	12	11	0.03	.87
No strategy/ knowledge	7	2	1.41	.24

Table 2. Percent of use of each strategy reported in the melody studies and timbre study and associated chi-square test results. Asterisks indicate $p < .05$.

4. DISCUSSION

Results showed that overall, most participants could report information about the strategy they believed they had used to remember sounds during a retention interval. Perhaps not surprisingly, an auditory-musical strategy was the most frequently reported strategy.

Some notable differences in strategy use were reported for remembering melodies versus timbre. Significantly more participants reported using a motor strategy to remember melodies as compared to timbre. This suggested that the pitch information in melodies may map more readily to motor responses (finger tapping, etc.). Research has indicated that higher pitch is associated with higher locations in space [31]. As such, people may be able to use motor encoding that represents pitch in physical space to supplement memory for melodies (see section 3.3.3). Previous research [17] has suggested that musicians may encode pitch using motor codes that involve production of the note on their instruments. The current study's results suggested that non-musicians might use more generic or rudimentary motor encoding strategies that link the pitch of tones to locations in space with motor programs.

Although participants reported high use of an auditory-musical encoding strategy for both melodies and timbre, this strategy was reported significantly more frequently in the melody studies. The sounds in the timbre study were abstract, and, although they were discriminable, some sounded perceptually similar. Perhaps these features discouraged or suppressed some participants' attempts to remember the timbral sounds as sounds per se and thus prompted them to attempt different encoding strategies.

Two strategies—verbal labeling and an associative strategy—were reported significantly more frequently among participants in the timbre study as compared to the melody study. Both strategies involve a version of elaboration—linking the abstract timbral stimuli to other information (by verbal labelling or associating the abstract sounds with existing knowledge) in an effort to aid memory. Elaborative strategies can aid in memory tasks (see [32]), especially when the bottom-up stimulus has no inherent meaning, as was the case here.

Taken together, the data presented here suggest that people use a robust variety of strategies to remember nonverbal sounds, and the most frequently used strategy was an auditory strategy that attempted to remember the sounds as sounds per se. Our findings suggested that abstract sounds may be more difficult than melodies to remember based on their acoustic properties, thus some people may try to encode them by linking them with information already present in their memory repertoire (e.g., verbal labels or names, other known sounds, etc.).

4.1. Limitations

Our approach in this study has limitations that should be considered when interpreting our findings. First, our data were limited to those strategies that participants chose to report. A participant's failure to mention using a particular strategy does not ensure that the participant did not use that strategy; instead, she may have used the strategy but chosen not to report it. The open-ended format of the question allowed participants to type as little as they wished. Some participants may have offered incomplete reports regarding their strategy use. Since the strategy question was asked at the end of the experiment, it also required participants to remember their strategies. It is possible that some participants had forgotten important aspects of their encoding strategies by the end of the study.

Further, participants can only report accurately on their strategy use to the extent that their memory processes were consciously accessible to them. Although the assumption that encoding strategies are accessible, reportable, and under voluntary command has been accepted in various research paradigms in cognitive psychology (e.g., [33], [34]), this assumption has also been challenged as an invalid appeal to introspection [35]. Our position on this matter lies somewhere in between these two perspectives. We acknowledge that the strategies reported by our participants cannot necessarily be assumed to accurately capture their actual mental processes. We also suspect, however, that in many cases the subjective reports do provide some information about those mental processes and point to interesting areas for future research with objective, quantitative approaches to strategy measurement. Further, we find the phenomenon of subjective impressions of strategy use to be one that is of interest in and of itself. That is, the phenomenological aspect of perception—"*what it is like to perceive*" a stimulus—has garnered serious attention as a topic of inquiry unto itself (see [36], pp. 2), yet few reports have examined this topic with respect to memory for melodies and timbre.

4.2. Directions for Future Research

Research consistently has suggested that multiple strategies can be used to encode pitch information [17]–[19], [21]. Less is known about encoding strategies for timbre. Researchers (e.g., [22]) have suggested that timbre memory entails auditory-sensory coding that maintains some isomorphic and phenomenologically auditory (e.g., rather than motor) representation of the timbral features of sounds.

Still, little is known about the functionality of cognitive mechanism(s) of auditory encoding strategies that preserve the acoustic characteristics of sounds. In particular, it is unclear how purely auditory (nonverbal, etc.) representations are rehearsed in working memory. Some researchers (e.g., [13],

[22], [23]) seem to have adopted the perspective that working memory for timbre entails a version of auditory sensory memory (see [6]) that is protracted in duration, though the mechanism that permits the representation to endure in working memory following perception has not been specified. More research is needed to establish how auditory representations are rehearsed in working memory.

5. CONCLUSIONS

Auditory displays and sonifications that use pitch to convey information may invoke different memory strategies than those that use abstract properties of sounds that must be discriminated based on timbre. Whether the use of different strategies ultimately will be an advantage or a hindrance to the end user will depend upon the particular use scenario (task constraints, etc.) and display design, but the deployment of auditory displays should proceed with awareness that different types of sounds may engage different cognitive processes in memory.

6. FOOTNOTES

1. The calculator used to compute the Chi-square tests is available online at:
https://www.medcalc.org/calc/comparison_of_proportions.php

7. REFERENCES

[1] G. Kramer, "An introduction to auditory display," in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, G. Kramer, Ed. Reading, MA: Addison Wesley, 1994, pp. 1–78.

[2] G. Kramer *et al.*, "The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display," 1999.

[3] J. H. Flowers, L. E. Whitwer, D. C. Grafel, and C. A. Kotan, "Sonification of daily weather records: Issues of perception, attention, and memory in design choices," in *Proceedings of the International Conference on Auditory Display,* Espoo, Finland, 2001, pp. 222–226.

[4] A. D. Baddeley, "Working memory: Theories, models, and controversies," *Annual Review of Psychology*, vol. 63, pp. 1–29, 2012.

[5] C. D. Wickens, "Multiple resources and performance prediction," *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, pp. 159–177, Jan. 2002.

[6] M. A. Nees, "Have We Forgotten Auditory Sensory Memory? Retention Intervals in Studies of Nonverbal Auditory Working Memory," *Frontiers in Psychology*, vol. 7, 2016.

[7] R. H. Logie and N. Cowan, "Perspectives on working memory: introduction to the special issue," *Memory & Cognition*, vol. 43, no. 3, pp. 315–324, 2015.

[8] A. Miyake and P. Shah, *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge University Press, 1999.

[9] A. D. Baddeley and R. H. Logie, "Auditory imagery and working memory," in *Auditory Imagery*, D. Reisberg, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, pp. 179–197.

[10] W. L. Berz, "Working Memory in Music: A Theoretical Model," *Music Perception: An Interdisciplinary Journal*, vol. 12, no. 3, pp. 353–364, Apr. 1995.

[11] K. Siedenburg and S. McAdams, "The role of long-term familiarity and attentional maintenance in short-term memory for timbre," *Memory*, vol. 25, no. 4, pp. 550-564.

[12] R. G. Crowder, "Auditory memory," in *Thinking in Sound: The Cognitive Psychology of Human Audition*, S. McAdams and E. Bigand, Eds. New York, NY, US: Clarendon Press/Oxford University Press, 1993, pp. 113–145.

[13] L. Demany and C. Semal, "The role of memory in auditory perception," in *Auditory Perception of Sound Sources*, Springer, 2008, pp. 77–113.

[14] T. A. Keller, N. Cowan, and J. S. Saults, "Can auditory memory for tone pitch be rehearsed?," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 21, no. 3, pp. 635–635, 1995.

[15] C. M. Macleod, E. B. Hunt, and N. N. Mathews, "Individual differences in the verification of sentence—picture relationships," *Journal of Verbal Learning and Verbal Behavior*, vol. 17, no. 5, pp. 493–507, Oct. 1978.

[16] N. N. Mathews, E. B. Hunt, and C. M. MacLeod, "Strategy choice and strategy training in sentence-picture verification," *Journal of Verbal Learning and Verbal Behavior*, vol. 19, no. 5, pp. 531–548, 1980.

[17] R. J. Zatorre and C. Beckett, "Multiple coding strategies in the retention of musical tones by possessors of absolute pitch," *Memory & Cognition*, vol. 17, no. 5, pp. 582–589, 1989.

[18] M. Mikumo, "Multi-encoding for pitch information of tone sequences," *Japanese Psychological Research*, vol. 39, no. 4, pp. 300–311, 1997.

[19] M. A. Nees and B. N. Walker, "Flexibility of Working Memory Encoding in a Sentence-Picture-Sound Verification Task," *Journal of Cognitive Psychology*, vol. 25, no. 7, pp. 800–807, 2013.

[20] M. A. Nees and K. Best, "Modality and encoding strategy effects on a verification task with accelerated speech, visual text, and tones," in *Proceedings of the International Conference on Auditory Display*, Lodz, Poland, 2013, pp. 267–274.

[21] M. A. Nees and B. N. Walker, "Mental scanning of sonifications reveals flexible encoding of nonspeech sounds and a universal per-item scanning cost," *Acta Psychologica*, vol. 137, no. 3, pp. 309–317, 2011.

[22] R. G. Crowder, "Imagery for musical timbre," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 15, no. 3, pp. 472–478, 1989.

[23] J. L. Golubock and P. Janata, "Keeping timbre in mind: Working memory for complex sounds that can't be verbalized," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 39, no. 2, pp. 399–412, 2013.

[24] M. A. Nees and B. N. Walker, "Encoding and representation of information in auditory graphs: Descriptive reports of listener strategies for understanding data," in *Proceedings of the International Conference on Auditory Display,* Paris, France, 2008.

[25] M. A. Nees and B. N. Walker, "Listener, task, and auditory graph: Toward a conceptual model of auditory graph comprehension," in *Proceedings of the International Conference on Auditory Display,* Montreal, Canada, 2007, pp. 266–273.

[26] M. A. Nees, Corrini, Ellen, Leong, Peri, and Harris, Joanna, "Maintenance of Memory for Melodies:

Articulation or Attentional Refreshing?," *Psychonomic Bulletin & Review*, In Press.

[27] M. A. Nees, Leong, Peri, and Harris, Joanna, "Mechanisms of working memory for timbre," Unpublished data.

[28] Z. A. Schendel and C. Palmer, "Suppression effects on musical and verbal memory," *Memory & Cognition*, vol. 35, no. 4, pp. 640–650, Jun. 2007.

[29] J. T. E. Richardson, "The analysis of 2 × 2 contingency tables—Yet again," *Statistics in Medicine*, vol. 30, no. 8, pp. 890–890, Apr. 2011.

[30] I. Campbell, "Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations," *Statistics in Medicine*, vol. 26, no. 19, pp. 3661–3675, Aug. 2007.

[31] E. Ben-Artzi and L. E. Marks, "Visual-auditory interaction in speeded classification: Role of stimulus difference," *Perception & Psychophysics*, vol. 57, no. 8, pp. 1151–1162, 1995.

[32] F. I. M. Craik and R. S. Lockhart, "Levels of processing: A framework for memory research," *Journal of Verbal Learning and Verbal Behavior*, vol. 11, no. 6, pp. 671–684, Dec. 1972.

[33] J. Marquer and M. Pereira, "Reaction times in the study of strategies in sentence–picture verification: A reconsideration," *The Quarterly Journal of Experimental Psychology Section A*, vol. 42, no. 1, pp. 147–168, 1990.

[34] S. M. Kosslyn, *Image and Mind*. Cambridge, MA: Harvard University Press, 1980.

[35] J. S. B. Evans, "Verbal reports of cognitive strategies: A note on Marquer and Pereira," *The Quarterly Journal of Experimental Psychology*, vol. 42, no. 1, pp. 169–170, 1990.

[36] W. Fish and others, *Philosophy of Perception: A Contemporary Introduction*. Routledge, 2010.

8. ACKNOWLEDGMENTS