
Examining Absolute and Relative Pitch as Automatic Processes via Auditory-Visual Stroop Test

Miren Edelstein

University of California, Berkeley

Relative pitch is the ability to identify intervals between pitches, whereas absolute pitch is the less common ability to identify and reproduce musical pitches without reference to other pitches. This study utilized an auditory-visual Stroop test to investigate the automaticity of absolute and relative pitch. Absolute and relative pitch possessors simultaneously heard a piano tone and viewed a music note in treble clef on a computer screen. The piano tone and viewed note were either congruent or incongruent with one another, and participants were asked to vocally identify either the visual or auditory stimuli. I hypothesized that absolute pitch possessors would exhibit slower reaction times than relative pitch possessors when identifying notes during incongruent as opposed to congruent pitch and note trials, which would suggest that absolute pitch possessors are unable to suppress pitch labeling and therefore experience cognitive interference. Contrary to my predictions, however, I found only a main effect of congruency, such that both absolute and relative pitch possessors were equally slowed down by incongruent pitch and note trials. These results suggest that relative pitch possessors may be able to acquire a sense of absolute pitch when exposed to multiple instances of congruent visual and auditory stimuli. In addition, these findings demonstrate that cognitive interference can occur cross-modally between two different sensory pathways such as audition and vision, and that this interference is not limited to individuals who have internalized and absolute (i.e., one to one) relationships between sounds and images.

Absolute pitch is the rare ability to identify or reproduce musical pitches without reference to other pitches. In general, it is a coveted ability amongst musicians; however, there is evidence that absolute pitch possessors tend to automatically rely on the ability even when it is not useful, which can hinder musical processing (Miyazaki, 2002). Miyazaki (2002) presented participants with a task designed to assess relative pitch, or the ability to identify intervals between notes. In this task, absolute pitch possessors appeared unable to suppress their absolute pitch, even though it detrimentally affected their performance on the task. Absolute pitch may therefore be an automatic process that can result in cognitive interference, a phenomenon in which intruding thoughts hinder performance on a task. The current research examined absolute pitch possessors to further explore the nature of cognitive interference.

The classic experiment documenting the phenomenon of cognitive interference is the Stroop test

(Stroop, 1935). Participants were shown color words printed in either congruent or incongruent colors (e.g., the word “red” printed in either red or blue). When asked to name the printed color of a color word that was incongruent, participants exhibited slower reaction times. This suggests that participants could not suppress reading the color word when asked to name the printed color, thereby demonstrating cognitive interference. Researchers have since found evidence for cognitive interference in other modalities of perception such as audition, vision and musical processing (Stewart, 2005; Miyazaki, 2004; Cohen & Martin, 1975; Cowan & Barton, 1987; Green & Barber, 1981; Leboe &

Author Note: This research was conducted under Frederic Theunissen, PhD. Please address correspondence regarding this article to Miren Edelstein. E-mail: miren.edelstein@gmail.com.

Mondor, 2007; Most, Sorber & Cunningham, 2007; Ward, Huckstep & Tsakanikos 2006). For instance, Stewart (2005) found that reading musical notation may be as automatic for trained musicians as reading printed words is for literate individuals.

Absolute pitch may similarly be automatic and result in cognitive interference. Miyazaki (2004) presented participants with tones that were sung as solfeggi syllables (e.g., *do, re, mi*) in varying pitches (e.g., C, D, E), and asked them to simply repeat the solfeggi syllables. Solfeggi syllables and pitches were either congruent (e.g., *do* sung as C, *re* sung as D) or incongruent (e.g., *do* sung as D, *re* sung as C). Absolute pitch possessors were significantly slower than relative pitch possessors to repeat the solfeggi syllables of the tones when pitches and syllables were incongruent, because the automatic identification of syllable with pitch among absolute pitch possessors hindered their ability to name the incongruent solfeggi syllables (Miyazaki, 2004).

One limitation of Miyazaki's (2004) study is that it examined absolute pitch and cognitive interference only with respect to the single sensory modality of audition. Prior research has indicated that the Stroop effect can occur cross-modally, such as between audition and vision (Leboe & Mondor, 2005). Leboe and Mondor (2005) presented participants with a series of tones that were high or low in frequency. Each tone was produced by a speaker in a physically high or low location. For each tone, participants were asked to judge whether the tone had a high or low frequency, or whether the tone was produced from a high or low physical location. Participants were significantly slower to judge the pitch or location of tones when pitch and location were incongruent (e.g., when a low-frequency tone was produced from a physically high location). Conflicting information from the auditory and visual modalities resulted in cognitive interference, which suggests that cognitive interference can occur cross-modally. The existence of cross-modal cognitive interference indicates that separate sensory pathways of the brain can influence one another, implying that the senses are more interconnected than researchers may previously have thought.

However, we propose that some people may be more likely than others to exhibit cross-modal cognitive interference. One study that examined tone-color synesthesia, a phenomenon in which people automatically associate musical tones with colors, found that synesthetes experienced cross-modal interference when presented with incongruent tone and color instances while controls did not

(Ward, Huckstep & Tsakanikos, 2006). In a way, synesthetes are similar to absolute pitch possessors because they both possess uncommon, automatic associations. Specifically, because absolute pitch possessors possess automatic associations between aural pitches and visual musical notations, we suggest that absolute pitch possessors may be particularly likely to experience cross-modal cognitive interference across audition and vision. In contrast, relative pitch possessors do not possess automatic associations between pitch and musical notations and should therefore be less likely to experience cross-modal cognitive interference.

The current study sought to examine differences in the way absolute and relative pitch possessors exhibit cross-modal, cognitive interference. Absolute and relative pitch possessors simultaneously heard a piano tone and viewed a congruent or incongruent music note in treble clef. Half of all trials were preceded by reference tones, which were provided both visually and aurally in order to help relative pitch possessors identify piano tones. We hypothesized that both absolute and relative pitch possessors would be slower to identify piano tones during incongruent as opposed to congruent piano tone and note pairings. However, we also hypothesized that only absolute pitch possessors would be slower to identify music notes during incongruent as opposed to congruent piano tone and note pairings, as this would be indicative of an inability to suppress pitch naming, leading to cross-modal cognitive interference.

METHOD

Participants. Thirty participants (15 males and 15 females; mean age = 21, range: 18-30) were recruited from the music department of a large public university in exchange for \$10. Participants were included in the study only if they: self-identified with having either absolute or relative pitch; were capable of fluently reading musical notation in treble clef; and did not have hearing impairments. Fifteen participants (7 males and 8 females; average amount of musical training = 15.5 years) personally identified with having absolute pitch, whereas the remaining fifteen participants (8 males and 7 females; average amount of musical training = 14.1 years) identified only with having relative pitch.

Procedure. The experiment was run on a Dell® desktop computer on the university campus. For

the purpose of the screening tests, a USB keyboard was altered with stickers so that specific keys corresponded with certain pitches and note names. A pair of Plantronics® headsets with a built in microphone was used to present auditory stimuli as well as record reaction times of vocal responses in milliseconds. The experiment was programmed and conducted on Presentation® (version 14.8). All auditory stimuli consisted of computerized piano tones, which were played at a comfortable volume level that was adjusted according to each participant's preference. These tones were tuned to an A of 440 Hz, the standard baseline for musical pitches, and spanned two octaves, ranging from C4 (also commonly known as middle C) to C6. Visual stimuli were presented on the computer screen as dotted quarter notes on a treble clef. All auditory and visual stimuli were created in the composition program Finale®.

After obtaining written consent, the experimenter administered the absolute pitch screening test, which consisted of 36 trials. For each trial, participants heard a piano tone and were asked to identify the note. Target tones were within an octave of each other, ascending or descending, and there was a 1000ms delay between each trial. There was no time limit placed on target tone identification. This test was administered to all participants in order to ensure consistency in the paradigm and to detect the presence of non-reported absolute pitch in relative pitch possessors.

The experimenter then administered the relative pitch screening test, which consisted of 36 trials. For each trial, participants heard a random reference note for 3000ms, which was followed by a delay of 1000ms and then the target piano tone. Reference notes were presented both visually (as a dotted quarter note on a treble clef) and aurally (as a computerized piano tone), whereas target tones were presented only aurally. The target piano tones were always within an octave of the reference note, ascending or descending. Participants were asked to identify the target tones using their keyboards. There was a 1000ms delay between each trial, and there was no time limit placed on target tone identification.

Next, the experimenter administered either version A or B of the auditory-visual Stroop test. Versions alternated between every other participant. The auditory-visual Stroop test consisted of two phases, which contained 86 trials each and were counterbalanced. Prior to each phase, 5 practice trials were presented to familiarize participants with the task. During each trial, participants simul-

taneously heard a computerized piano tone and saw a dotted quarter note shown on a treble clef. The piano tone and dotted quarter notes were congruent for half of all trials, and incongruent for the other half. There was a 1000ms delay between each trial. In phase 1, participants were asked to vocally identify the piano tone they heard during each trial. In phase 2, participants were asked to vocally identify the dotted quarter note they saw during each trial. Reference tones were randomly inserted prior to half of all trials in order to help relative pitch possessors identify the piano tones in phase 1. Reference tones were presented for 3000ms, and the actual trial was presented following a 1000ms delay. During actual trials, participants had a 6000ms time limit to make their vocal response.

Finally, participants provided basic demographic information (e.g., age, gender, ethnicity) and were fully debriefed.

RESULTS

Absolute Pitch Screening Test. All self-identified absolute pitch possessors scored above the recommended cutoff score of 80% for absolute pitch possessors (Dooley & Deutsch, 2010) on the absolute pitch screening test ($M = 96.40$, $SD = 5.34$, range: 83-100). Scores of relative pitch possessors on the absolute pitch screening test varied greatly ($M = 24.15\%$, $SD = 15.88\%$, range: 0-53%). Chance performance on the absolute pitch screening test was calculated at 8.33%, equivalent to a 1 in 12 chance of responding correctly on each trial across 36 trials (Dooley & Deutsch, 2010). On average, relative pitch possessors did significantly better than chance on the test ($t(14) = 3.85$, $p < .001$) suggesting that some may possess borderline absolute pitch and may be able to consistently identify certain pitches although not all.

Relative Pitch Screening Test. Absolute pitch possessors generally scored well on the relative pitch screening test ($M = 96.60\%$, $SD = 9.27\%$, range: 64-100%), but the score they received was not included in the final data analysis. Relative pitch possessors showed a marked improvement on the relative pitch screening test from the absolute pitch screening test ($M = 66.47\%$, $SD = 15.40\%$, range: 42-92%). This improvement in performance was significant ($t(28) = 7.41$, $p < .0001$), indicating that the inclusion of reference tones was beneficial to

the performance of relative pitch possessors and that they did indeed demonstrate the ability of relative pitch.

Auditory-Visual Stroop Test. As a disclaimer, it is important to note that the phase 1 trials of relative pitch possessors were not included in the final data set due to the task being too difficult for the major-

ity of participants to complete within the set time frame. Thus, a comparison between absolute and relative pitch possessors could only be made when examining phase 2 trials. We hypothesized that absolute pitch possessors should exhibit more of a cross-modal Stroop effect than relative pitch possessors, exhibiting slower reaction times in incongruent trials regardless of phase.

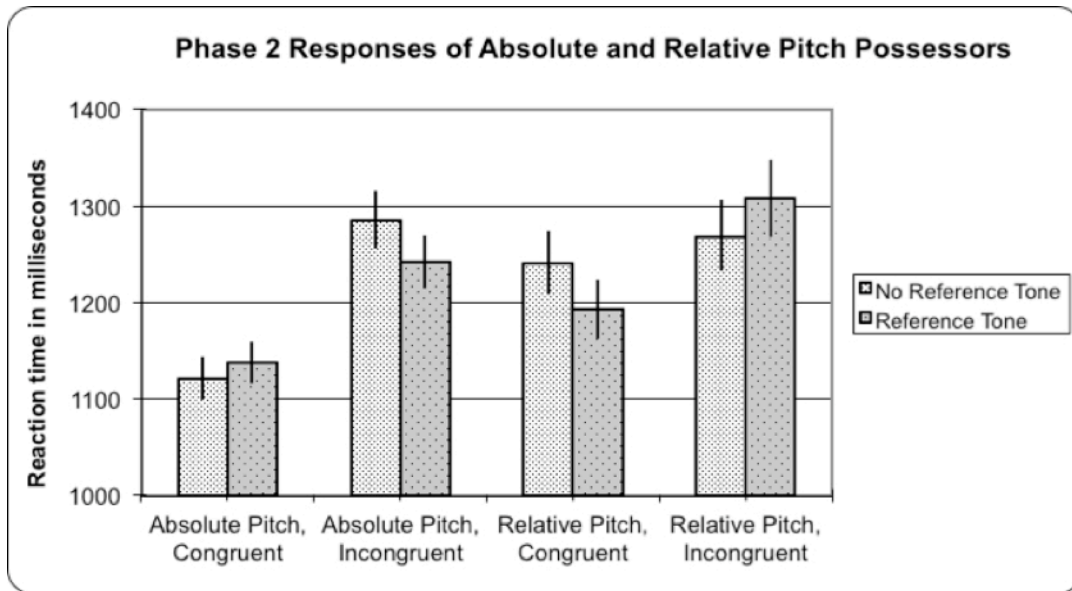


FIGURE 1. Mean reaction times of Phase 2 trials. Shows mean reaction times of congruent and incongruent trials, with and without being preceded by a reference tone, of absolute and relative pitch possessors.

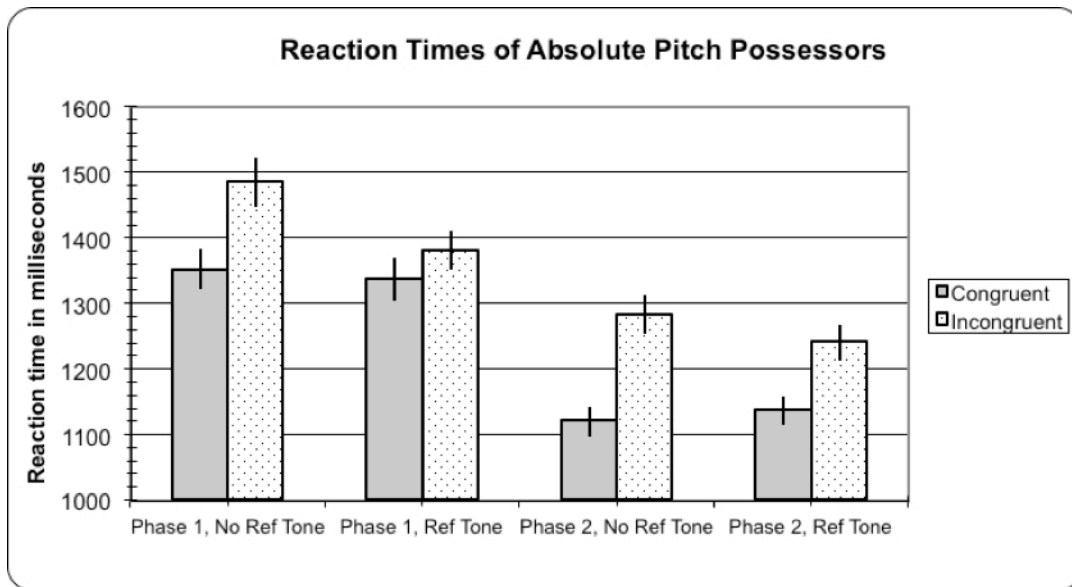


FIGURE 2. Mean reaction times of congruent and incongruent trials, with and without being preceded by a reference tone, of absolute pitch possessors in both Phases 1 and 2.

Comparing absolute and relative pitch possessors in phase 2 (see Figure 1). We conducted a 2 (Subject Type: absolute pitch possessor or relative pitch possessor) \times 2 (Congruency: congruent or incongruent) \times 2 (Reference Tone: present or absent) repeated-measures ANOVA on reaction time. In contrast to our hypothesis that incongruent trials would increase the reaction times of absolute pitch possessors but not relative pitch possessors, we found only a main effect of Congruency, $F(1, 28) = 12.22$, $p = .002$, $\eta_p^2 = .30$, such that participants responded 97ms faster during congruent trials than during incongruent trials. No other main effects or interactions were significant, $ps > .3$.

Comparing phases 1 and 2 in absolute pitch possessors (see Figure 2). We conducted a 2 (Phase: 1 or 2) \times 2 (Congruency: congruent or incongruent) \times 2 (Reference Tone: present or absent) repeated-measures ANOVA on reaction time. In support of our hypotheses, there was a marginally significant main effect of Phase, $F(1, 14) = 4.23$, $p = .059$, $\eta_p^2 = .23$, such that absolute pitch possessors responded 189ms faster in phase 2 than in phase 1. There was a significant main effect of Congruency, $F(1, 14) = 13.30$, $p = .003$, $\eta_p^2 = .49$, such that absolute pitch possessors responded 104ms faster during congruent trials than during incongruent trials. Finally, there was a marginally significant main effect of reference tone, $F(1, 14) = 4.40$, $p = .055$, $\eta_p^2 = .24$, such that absolute pitch possessors responded 48ms faster when a reference tone was present than when it was absent. There were no significant interactions, $ps > .1$.

Phase 1 data in absolute pitch possessors. We conducted a 2 (Congruency: congruent or incongruent) \times 2 (Reference Tone: present or absent) repeated-measures ANOVA on reaction time. As predicted, there was a significant main effect of Congruency, $F(1, 14) = 7.33$, $p = .017$, $\eta_p^2 = .34$, such that absolute pitch possessors responded 94ms faster during congruent trials than during incongruent trials. There was also a marginally significant main effect of Reference Tone, $F(1, 14) = 4.00$, $p = .065$, $\eta_p^2 = .22$, such that absolute pitch possessors responded 68ms faster when a reference tone was present than when it was absent. There was no significant interaction, $p = .165$.

DISCUSSION

Because we unfortunately could not include phase 1 trials from relative pitch possessors in our final dataset, we were unable to validate whether relative pitch possessors would be slower to identify piano tones during incongruent as opposed to congruent piano tone and note pairings. However, we also hypothesized that only absolute pitch possessors, as opposed to relative pitch possessors, would be slower to identify music notes during incongruent as opposed to congruent piano tone and note pairings. Although we did find that absolute pitch possessors were slower to identify music notes during incongruent as opposed to congruent trials, we surprisingly found that relative pitch possessors exhibited the very same effect. This unexpected finding indicates that although relative pitch possessors cannot automatically identify pitches nearly as well as absolute pitch possessors, they may still be able to subconsciously perceive whether a pitch is congruent to a visually perceived note.

Contrary to Miyazaki's (2004) results, our findings suggest that absolute pitch possessors are not the only ones susceptible to cognitive interference induced by a musical Stroop effect. Since all relative pitch possessors in the present study had played musical instruments for many years, it is possible that their extensive training may have given them the ability to subconsciously detect whether pitch and note pairings were correct. That there were no significant differences in reaction times between absolute and relative pitch possessors provides further evidence that the two groups may actually be more similar to each other than previously thought. Moreover, our findings reinforce previous findings that cross-modal cognitive interference can indeed occur, but in addition, they also suggest that it may occur subconsciously.

The fact that relative pitch possessors exhibited the same cognitive interference as absolute pitch possessors in phase 2 is particularly interesting. A learning effect due to the presence of reference tones is another possible contributor to the unexpected discrepancy in reaction times observed between congruent and incongruent trials in relative pitch possessors. When presented with reference tones, and therefore correct pitch and note pairings, relative pitch possessors may take them into account subconsciously. As a result, when incorrect pitch and note pairings occur, relative pitch possessors may exhibit slower reaction times because although they may not be able to pinpoint the exact

note that is being played, they are still able to sense that it is incorrect. This sense that the pairing is incorrect may be salient enough to hinder reaction times.

To assess just how much reference tones may have affected the performance of relative pitch possessors, a future study can be conducted without reference tones. This way, relative pitch possessors would not have any external confirmation of correct pitch and note pairings. If the same cognitive interference is still present even in the absence of reference tones, it would be safe to assume that relative pitch possessors have subconsciously retained associations between pitches and notes due to years of musical training. However, if the cognitive interference disappears with the reference tones, this would suggest that relative pitch possessors in the present study were actively acquiring a form of absolute pitch scaling from the provided reference tones throughout the duration of the study. This would have significant implications regarding the speed at which subconscious associations can be formed.

We note several limitations of the present study. Over the course of the experiment, it became apparent that the auditory-visual Stroop test was too difficult for most relative pitch possessors to comfortably complete within the given timeframe. This is likely due to the fact that the reference tones changed every time they were presented, which forced relative pitch possessors to constantly shift their tonal centers. If we had instead provided a reference tone that was constant throughout the experiment, such as C4 (also known as middle C), the relative pitch possessors may have performed better on the test.

In addition, although most relative pitch possessors found both the auditory-visual Stroop and relative pitch screening tests to be quite difficult, we note that four relative pitch possessors actually performed quite well on both the absolute and relative pitch screening tests, even scoring above 40% on the absolute pitch screening test. These four participants also received some of the highest scores on the relative pitch screening test. Thus, it may have been beneficial to create a third group of participants (borderline absolute pitch possessors), because these participants may have reduced our ability to detect significant differences between relative and absolute pitch possessors in our data.

Despite these limitations, the study still revealed several intriguing findings. Essentially, absolute pitch possessors can be susceptible to bidirectional, cross-modal cognitive interference and rela-

tive pitch possessors may either be rapidly acquiring subconscious associations of pitch and note, or retaining more pitch and note information than they are aware of. Ultimately, visual stimuli can affect the processing of auditory stimuli and vice versa. The findings from this study have shed light on and raised thought-provoking questions about how the musical mind processes and combines auditory and visual cues. We aim to further investigate the questions raised in this study through future research endeavors.

References

- Cohen, G., & Martin, M. (1975). Hemispheric differences in an auditory Stroop task. *Perception and Psychophysics*, 17, 79-83.
- Cowan, N., & Barton, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, 41, 393-401.
- Dooley, K. & Deutsch, D. (2010). Absolute pitch correlates with high performance on musical dictation. *Journal of the Acoustical Society of America*, 128, 890-893.
- Green, E. J., & Barber, P. J. (1981). An auditory Stroop effect with judgments of speaker gender. *Perception & Psychophysics*, 30, 459-466.
- Leboe, L. C., & Mondor, T. A. (2007). Item-specific congruency effects in nonverbal auditory Stroop. *Psychological Research*, 71, 568-575.
- Miyazaki, K. and Rakowski, A. (2002). Recognition of notated melodies by possessors and non possessors of absolute pitch. *Perception and Psychophysics*, 64, 1337-1345.
- Miyazaki, K. (2004). The Auditory Stroop Interference and the Irrelevant Speech/Pitch Effect: Absolute-Pitch Listeners Can't Suppress Pitch Labeling. *Paper presented at the 18th International Congress on Acoustics (ICA 2004)*, April 4-9, 2004 (Kyoto, Japan).
- Most, S. B., Sorber, A. V., & Cunningham, J. G. (2007). Auditory Stroop reveals implicit gender associations in adults and children. *Journal of Experimental Social Psychology*, 43, 287-294.
- Stewart, L. (2005). A Neurocognitive Approach to Music Reading. *Annals of the New York Academy of Sciences*, 1060, 377-386.
- Stroop, J. R. (1935). Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Ward, J., Huckstep, B., Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, 42, 264-280.