Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1152–1162
https://doi.org/10.54941/ahfe1008345

From Supervision to Self-Correction: TuneTrace, a Persistent Visual Intonation Companion for Practice Autonomy in Autistic Young Musicians

Lujing Yang

Shanghai American School, Shanghai, China

ABSTRACT

Instrumental practice requires learners to notice pitch deviations, decide what to change, and evaluate another attempt. Adult guidance can support this process, but feedback that depends entirely on adult interpretation may leave limited opportunities for learner-initiated decisions. This paper presents TuneTrace, a browser-based intonation companion designed to support practice autonomy in autistic young musicians through persistent, learner-interpretable visual feedback. The design links pitch deviations to notation, retains annotations for review, and lets learners adjust when and how much feedback they receive. Its rationale combines presuming competence, self-regulated musical learning, and research on augmented feedback. The public demonstration exposes instrument and tuning settings, measure looping, and three feedback schedules: Live Support, Measure Review, and Standard Notation with end-of-session review. An exploratory eight-week evaluation with six autistic musicians aged 11–15 describes fewer adult interventions, more reported unassisted practice time, and more unprompted re-attempts during use. Re-attempts are distinguished from successful pitch correction, and these descriptive observations do not establish causal effects or learning transfer. The contribution is a design account of persistent feedback as a resource for learner-controlled practice, together with feasibility observations and a focused agenda for technical, accessibility, and educational validation.

Keywords: Autism, Music education, Intonation feedback, Practice autonomy, Self-regulated learning, Accessible interaction

INTRODUCTION

Playing in tune involves more than producing the correct note. A musician must compare the sound with an intended pitch, recognize the direction of a deviation, and choose an appropriate adjustment. Practice therefore combines auditory perception, instrumental control, and self-evaluation. Augmented feedback can make errors easier to notice, but performance while feedback is present must be distinguished from learning that persists after its removal (Salmoni, Schmidt and Walter, 1984; Schmidt and Lee, 2011). In musical learning, goal setting, monitoring, and reflection are central to self-regulation (McPherson and Zimmerman, 2002).

Received September 26, 2026; Revised October 29, 2026; Accepted November 3, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

For autistic young musicians, the design challenge is to offer usable support without assuming either incapacity or a uniform pattern of strengths. Teachers and families provide valuable expertise and reassurance. However, when they identify every deviation and direct every response, learners may have fewer opportunities to interpret feedback and initiate their own next step. This is a design hypothesis about the organization of practice, not an established explanation of dependence among autistic students. Presuming competence invites educators and designers to make opportunities for agency available while remaining responsive to individual support needs (Biklen and Burke, 2006).

TuneTrace explores this opportunity by keeping pitch feedback available after the sound has ended. A learner can play a phrase, inspect its trace, select a passage to repeat, and decide whether to continue or seek help (Figure 1). Here, practice autonomy means meaningful control over goals, feedback, repetition, and help-seeking. It does not require practicing alone or permanently abandoning an assistive tool. Persistence is intended to make interpretation possible at the learner's pace; whether it achieves that aim requires direct evaluation.

The paper contributes a design rationale for persistent visual intonation feedback, an account of TuneTrace's interaction and calibration choices, and descriptive observations from a small formative evaluation. Three questions organize the work: how can feedback support learner-initiated practice decisions; how can its timing and visual density be adjusted without removing needed support; and what evidence would distinguish increased engagement with feedback from successful self-correction and transferable learning?

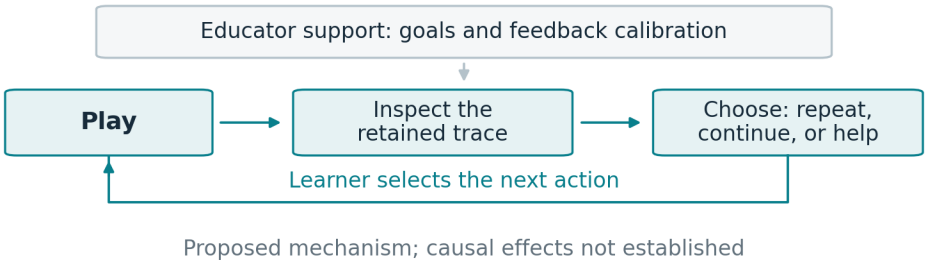


Figure 1: Proposed learner-controlled practice loop. The diagram expresses a design hypothesis, not a demonstrated causal effect of TuneTrace.

BACKGROUND AND RELATED WORK

Presuming Competence in Music Education

Neurodiversity perspectives challenge accounts that reduce autism to deficits and draw attention to interactions between people and environments (Kapp et al., 2013). Presuming competence is an educational commitment to accessible opportunities, rather than a prediction that every learner will perform a task without support (Biklen and Burke, 2006). Strengths-based music education similarly emphasizes adapting teaching to the individual

(Hammel and Hourigan, 2011). TuneTrace adopts this stance by treating feedback interpretation as a capability to support and investigate, while preserving the learner's right to request assistance.

Autism, Pitch Perception, and Instrumental Learning

Studies of autistic participants have documented preserved or enhanced pitch memory, labelling, and aspects of auditory perception (Heaton, 2003; Mottron et al., 2006; Heaton, 2009). Music can offer meaningful opportunities to recognize abilities that other tasks overlook (Molnar-Szakacs and Heaton, 2012). These findings challenge blanket assumptions of musical incapacity, but they do not establish universal pitch strengths or independent instrumental self-correction. Hearing a difference, interpreting a visual trace, executing a motor adjustment, and choosing to repeat a phrase are distinct demands. Musical experience, sensory preferences, communication, and support needs must therefore inform design and evaluation.

Augmented Feedback and the Guidance Problem

The guidance literature distinguishes immediate performance benefits from retention and transfer, and shows that feedback frequency can affect motor learning (Salmoni, Schmidt and Walter, 1984; Winstein and Schmidt, 1990). The challenge-point framework further emphasizes matching information to learner skill and task difficulty (Guadagnoli and Lee, 2004). These findings motivate adjustable feedback, but do not prove that any particular visual fading interface will improve learning. TuneTrace separates three properties: when feedback appears, how much visual information it contains, and how long it remains available. Reducing cue density is not equivalent to reducing feedback frequency.

Visual Intonation Feedback and Accessible Music Technology

Prior intonation systems show both promise and variation. Pardue and McPherson (2019) found heterogeneous responses among beginning violinists, including distraction, whereas Blanco, Tassani and Ramirez (2021) reported transfer benefits from visual feedback with knowledge of results in adult novice pitch-matching tasks. Neither establishes effects for autistic young musicians using persistent score annotations. Accessible musical-interface research also emphasizes customization and participation in design (Frid, 2019). CAST's guidelines support learner choice and customizable representation (CAST, 2024); W3C guidance requires information conveyed by color to be available through other visual means (W3C, 2023). TuneTrace brings these concerns together around the learner's opportunity to inspect and act on a retained practice trace.

DESIGN PRINCIPLES

Three principles guide TuneTrace's design (Table 1). They are design commitments to be assessed with learners, rather than established properties of autistic users or guarantees of accessibility.

P1 — Stable meaning within the feedback display. Within an intonation annotation, a color should carry a consistent meaning across notes and repetitions. The annotation needs an explicit legend and a stable pitch reference so that learners can relate successive attempts. This principle applies to performance feedback; it does not imply that every interface color, including navigation or branding, represents intonation.

P2 — Redundant visual encoding. Color is supplemented by direction markers, outline differences, and optional cent values. These are multiple visual encodings, not separate sensory modalities. They provide alternatives to hue discrimination (Okabe and Ito, 2008; W3C, 2023), but low-vision access, contrast, symbol comprehension, and sensory comfort still require evaluation. Labels should describe the relation to a target rather than judge the learner. A visually simple display may help some users and constrain others, making customization essential.

P3 — Learner-controlled scaffolding. Learners should be able to receive immediate feedback, defer inspection, or hide cues while they listen and judge a phrase themselves. Timing and cue density can be adjusted with an educator, without treating reduced support as mandatory progress. The educational hypothesis is that delaying inspection can create space for self-estimation; the corresponding risk is that insufficient feedback makes practice frustrating. Both possibilities must be assessed rather than inferred from a selected interface mode.

Table 1: Design principles and the questions they create for evaluation.

Principle	Design Realization	Evaluation Question
P1: Stable meaning	Consistent pitch reference and annotation legend	Can learners interpret the same cue across attempts?
P2: Redundant visual encoding	Hue plus direction, outline, and optional cent values	Are cues legible and understandable across visual needs?
P3: Learner-controlled scaffolding	Adjust feedback timing and visual density separately	Does delaying review support self-estimation without frustration?

THE TUNETRACE SYSTEM

Practice Setup and Public Demonstration

The public demonstration is available at https://tune-trace.replit.app/. Its setup screen offers Flute in C and B-flat Trumpet, an A4 reference from 438 to 444 Hz, tempo adjustment, and acceptance tolerances of ±5, ±10, or ±15 cents. The inspected default was A4 = 440 Hz and ±10 cents. The practice view includes notation, a measure counter, Loop Measure, Start Listening, and Finish Session controls (Figure 3). The flute exercise menu includes long tones, chromatic thirds, an arpeggio study, a simple melody, and high-register entry. These visible controls establish the demo's interaction scope; they do not by themselves validate pitch accuracy, latency, or data handling.

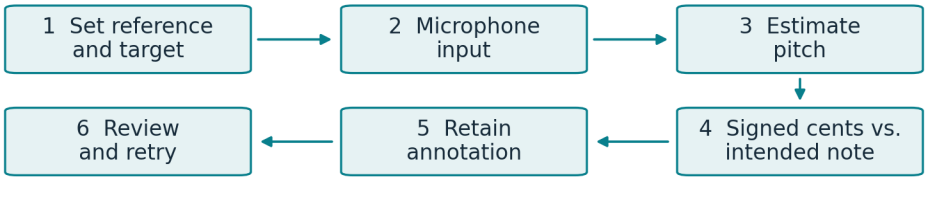


Figure 2: Conceptual workflow linking pitch estimation, persistent annotation, and learner review. Processing and privacy properties require technical validation.

Pitch Estimation and Interpretable Deviation

The intended processing loop captures a monophonic performance, estimates fundamental frequency, relates it to the intended score note, and retains an annotation for review (Figure 2). YIN provides a relevant method for fundamental-frequency estimation (de Cheveigné and Kawahara, 2002), but the deployed algorithm, frame settings, and latency require a reproducible implementation record. For a valid estimate f_i, and the sounding target frequency $\hat{f}_i$,, the signed deviation is:

$$d_i = 1200 \cdot \log_2\left(\frac{f_i}{\hat{f}_i}\right) \tag{1}$$

Positive d_i indicates a sharp pitch and negative d_i a flat pitch; 100 cents equal one equal-tempered semitone. The target must account for concert pitch and instrument transposition. Comparing only with the nearest chromatic pitch would miss wrong-note errors. Table 2 records the three-band design scheme; its 15-cent boundary is distinct from the demo's selectable acceptance tolerance. These are pedagogical settings, not universal perceptual limits. A validated pipeline must separate stable notes from attacks, silence, background sound, and tracking failures; uncertain estimates should remain unscored.

Table 2: Proposed three-band encoding, distinct from the demo tolerance setting.

Accuracy Band	Deviation |d|	Hue	Non-Color Cue
Within target	≤ 15 cents	Teal	Thin outline; check mark
Moderate deviation	> 15 to ≤ 35 cents	Amber	Medium outline; sharp/flat arrow
Larger deviation	> 35 cents	Vermilion	Heavy outline; sharp/flat arrow

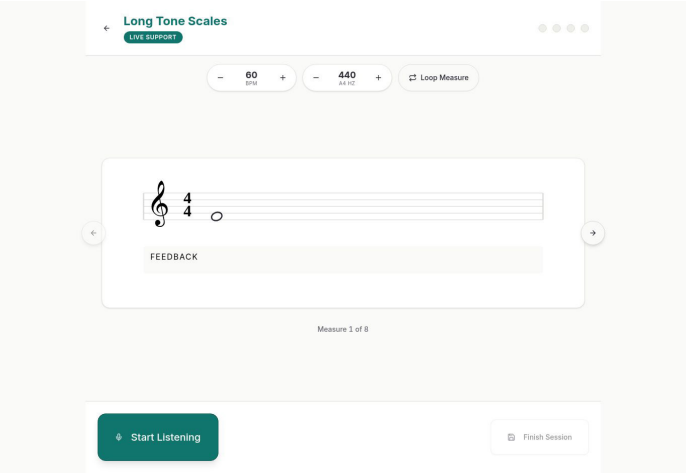


Figure 3: TuneTrace's public practice view before microphone capture. Tempo, tuning reference, measure looping, notation, and session controls are visible. This screenshot contains no participant performance data.

Persistence as a Resource for Reflection

The central design idea is to retain the relation between a performed note and its target after the note ends. Instead of requiring an immediate response to a moving tuner display, a learner can finish a phrase and then inspect a stable record. A retained trace can support decisions about what to repeat, but persistence alone does not identify the right instrumental adjustment or demonstrate understanding. Concurrent feedback can remain available for later review; it should not be called terminal feedback merely because it persists. Repeated-attempt comparison also requires an explicit rule for retaining or replacing annotations. The public demo's supplied exercises do not establish arbitrary MusicXML import or robust score following through stops and restarts.

Feedback Timing and Cue Density

The public demo exposes three schedules (Figure 4). Live Support presents annotations during held notes; Measure Review defers them until a measure is completed; Standard Notation hides support during practice and offers review at the end. Separately, the four-level design specification describes full color, markers, and cent labels (L1); color and markers (L2); neutral markers (L3); and plain notation with later reveal (L4). These levels change cue density, whereas the demo schedules change timing. They are not interchangeable experimental conditions. In either hidden-feedback arrangement, a later reveal implies that measurement continues even though cues are not displayed.

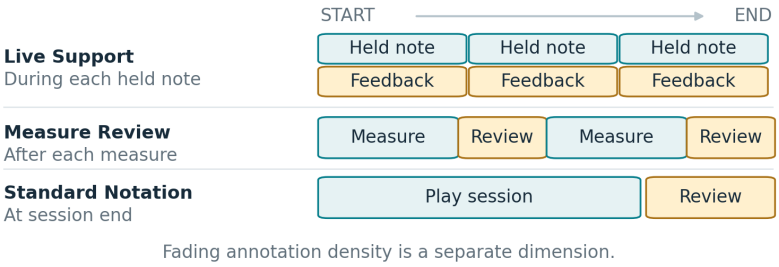


Figure 4: Feedback timing options named in the public demo. The four-level cue-density specification is a separate design dimension, not a fourth demo mode.

Calibration and Learner Control of Records

Calibration lets an educator help select achievable targets while leaving practice decisions with the learner. Privacy should support this relationship through data minimization, understandable retention controls, and learner choice over sharing. The public demo includes a progress page that invites users to complete sessions to populate charts. Consequently, an absolute claim that all traces disappear when the practice tab closes is not established. Local pitch processing, if implemented, would also not demonstrate the absence of telemetry or stored summaries. Audio handling, session retention, export, deletion, and adult access need separate technical verification; research observation records require their own consent and governance.

FORMATIVE FIELD EVALUATION

Setting and Participants

The formative evaluation involved six autistic students aged 11–15 in an inclusive middle-school concert band and an eight-week period of home practice. Participants volunteered with their families. Table 3 summarizes the participant profiles and prior practice arrangements. Independent practice was limited and varied: three profiles described no independent practice and three described occasional practice with nearby or periodic adult support. The evaluation explored use and design tensions in one setting; it did not include a comparison group or establish a treatment effect.

Table 3: Participant overview (pseudonyms).

Pseudonym	Age	Instrument	Prior Independent Practice
Alex	12	Flute	None; parent listened throughout
Ben	11	Trumpet	Occasional, parent nearby
Casey	14	Clarinet	None; parent corrected errors
Devon	13	Alto saxophone	Occasional, teacher-checked weekly
Emery	15	Trombone	None; parent listened throughout
Finley	12	Euphonium	Occasional, parent nearby

Measures and Procedure

The observation schedule specified one 30-minute home session per participant per week, with a non-intervening researcher logging adult-initiated corrections or prompts, minutes classified as unassisted practice, and the proportion of out-of-tune notes re-attempted without adult prompting. Out-of-tune notes were defined using $|d| > 15$ cents and checked against the session trace. A four-item, five-point confidence questionnaire was administered in weeks 1 and 8; it is treated here as an exploratory self-report measure. Learners began with L1 or L2 and could adjust cue density. Week 1 was the first observed week of use, not a documented no-tool baseline.

Descriptive Observations and Their Interpretation

Reported group means for adult interventions decreased from 11 per observed session in week 1 to 3 in week 8. Reported unassisted practice increased from 4 to 18 minutes of a 30-minute session, and the unprompted re-attempt rate increased from 18% to 68% (Figure 5). A re-attempt indicates renewed engagement with a note, not necessarily a correct diagnosis or an improved subsequent pitch. Likewise, unassisted minutes are a context-specific observation category, not proof that the learner was alone or experienced greater autonomy. These endpoints describe changes during use; they cannot isolate TuneTrace from repeated practice, changing adult behavior, or observer effects.

Confidence ratings increased across the four reported items; the mean for "I can practise without help" rose from 1.8 to 4.0. Two observations illustrate possible uses of the trace. Casey (clarinet) noticed a recurring flat indication, adjusted her playing, and saw the annotation change to the target band. Emery (trombone) used L4 to state a judgment before revealing the trace. These accounts suggest opportunities for noticing and self-estimation, but agreement with the same system that supplies feedback is not an independent test of auditory learning. They are illustrative cases, not evidence that every participant followed the same trajectory.

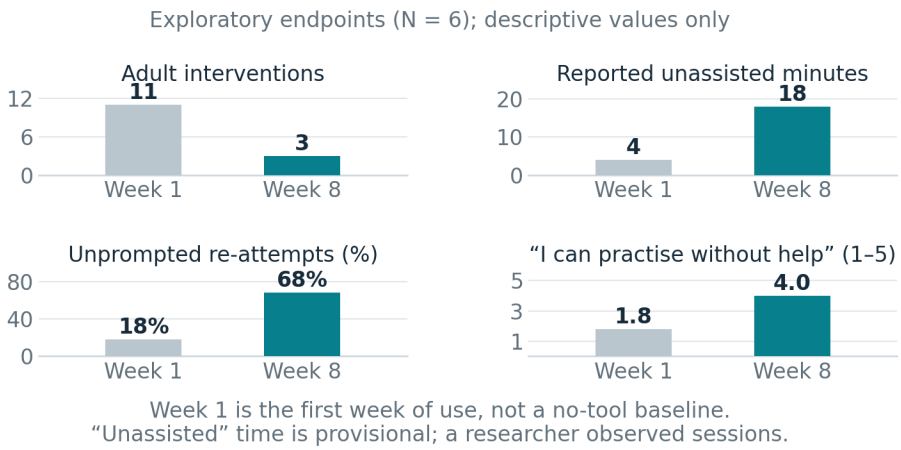


Figure 5: Reported week-1 and week-8 endpoints (N = 6). Re-attempts are not successful corrections. Week 1 represents use of the tool; uncertainty intervals and causal effects cannot be inferred from these summaries.

Two design tensions emerged in the field account. Sound from another person practicing nearby sometimes contaminated pitch tracking, indicating the need to test source discrimination and uncertain-estimate handling. One learner repeatedly pursued a favorable color on isolated notes at the expense of phrase continuity; phrase-level goals were introduced in response. These observations raise a shared concern: a feedback trace may support reflection, but it may also redirect attention toward the system's classifications rather than musical intention.

DISCUSSION

TuneTrace's strongest design proposition concerns access to practice decisions. A retained trace can give a learner material for deciding where to focus, whether to repeat, and when to seek help. This extends the monitoring and reflection processes described in self-regulated musical learning (McPherson and Zimmerman, 2002). The formative observations are consistent with greater learner-initiated activity during use, but fewer adult prompts cannot alone establish that authority shifted to the learner. Adults may simply have changed their behavior, and a learner may follow system cues without understanding them. Interviews and observation of decisions are needed alongside counts.

Persistent, learner-interpretable feedback is therefore proposed as a candidate design pattern with three requirements: a trace remains available long enough to inspect; its meaning can be understood without continual expert translation; and the learner has meaningful control over its use and sharing. Each requirement needs evidence. Comprehension tasks should test whether learners can distinguish direction, magnitude, and uncertainty; observation should establish whether that understanding informs a chosen action. Persistence within a phrase, across attempts, and across sessions are different retention choices and should not be conflated.

The shift from supervision to self-correction should preserve supportive relationships. A learner may autonomously choose a teacher's help, continue using visual cues, or prefer shared practice. Delayed review can create an opportunity to compare a personal judgment with the trace, but should remain reversible when it increases frustration. Research on augmented feedback motivates testing retention and transfer, not assuming them (Winstein and Schmidt, 1990). For music education, successful correction also involves phrasing, tone, and an appropriate musical pitch reference; optimizing every note's displayed color is an incomplete practice goal.

LIMITATIONS AND FUTURE WORK

This evaluation is limited by its small sample, one setting, lack of a comparison condition, and descriptive summaries. Individual trajectories, observation completeness, coding reliability, and measurement uncertainty are needed for stronger interpretation. The design combines persistence, adjustable feedback, calibration, and changes in practice relationships, so their contributions cannot be separated. The public demo's two instrument choices also do not establish support for all six instruments in the field account.

Technical priorities include pitch and note-alignment benchmarks under realistic noise, documented versioning, and verification of retention behavior. A future study should compare transient and persistent feedback with matched calibration, measure successful correction separately from re-attempts, and include no-cue retention and transfer tasks. Participatory work with autistic learners should examine comprehension, sensory comfort, desired support, and control over sharing. Such work is needed before making claims of efficacy, accessibility, or broad applicability.

CONCLUSION

TuneTrace investigates how persistent visual intonation feedback can make musical practice decisions more accessible to autistic young musicians. Its contribution is the combination of an inspectable trace, adjustable feedback, and an explicit focus on learner control. The formative observations describe fewer adult interventions and more unprompted re-attempts during use, while leaving successful correction and learning transfer unresolved. The next step is to test whether learners understand and use the trace to pursue their own musical goals, with support available on their terms.

ACKNOWLEDGMENT

The author thanks the participating students and families and the band director for their contribution to the formative work. The authors acknowledge the use of Grammarly for assistance with translation from Chinese to English during the preparation of this manuscript. The authors have reviewed and edited the final text and assume full responsibility for its content and accuracy.

REFERENCES

Biklen, D. and Burke, J. (2006) 'Presuming competence', Equity & Excellence in Education, 39(2), pp. 166–175. doi: 10.1080/10665680500540376.

Blanco, A.D., Tassani, S. and Ramirez, R. (2021) 'Effects of visual and auditory feedback in violin and singing voice pitch matching tasks', Frontiers in Psychology, 12, article 684693. doi: 10.3389/fpsyg.2021.684693.

CAST (2024) CAST Universal Design for Learning Guidelines version 3.0. Available at: https://udlguidelines.cast.org/ (Accessed: 26 September 2026).

de Cheveigné, A. and Kawahara, H. (2002) 'YIN, a fundamental frequency estimator for speech and music', Journal of the Acoustical Society of America, 111(4), pp. 1917–1930. doi: 10.1121/1.1458024.

Frid, E. (2019) 'Accessible digital musical instruments—A review of musical interfaces in inclusive music practice', Multimodal Technologies and Interaction, 3(3), article 57. doi: 10.3390/mti3030057.

Guadagnoli, M.A. and Lee, T.D. (2004) 'Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning', Journal of Motor Behavior, 36(2), pp. 212–224.

Hammel, A. and Hourigan, R. (2011) Teaching music to students with special needs: A label-free approach. New York: Oxford University Press.

Heaton, P. (2003) 'Pitch memory, labelling and disembedding in autism', Journal of Child Psychology and Psychiatry, 44(4), pp. 543–551. doi: 10.1111/1469-7610.00143.

Heaton, P. (2009) 'Assessing musical skills in autistic children who are not savants', Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1522), pp. 1443–1447. doi: 10.1098/rstb.2008.0327.

Kapp, S.K., Gillespie-Lynch, K., Sherman, L.E. and Hutman, T. (2013) 'Deficit, difference, or both? Autism and neurodiversity', Developmental Psychology, 49(1), pp. 59–71. doi: 10.1037/a0028353.

McPherson, G.E. and Zimmerman, B.J. (2002) 'Self-regulation of musical learning: A social cognitive perspective', in Colwell, R. and Richardson, C. (eds.) The new handbook of research on music teaching and learning. New York: Oxford University Press, pp. 327–347.

Molnar-Szakacs, I. and Heaton, P. (2012) 'Music: A unique window into the world of autism', Annals of the New York Academy of Sciences, 1252(1), pp. 318–324.

Mottron, L., Dawson, M., Soulières, I., Hubert, B. and Burack, J. (2006) 'Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception', Journal of Autism and Developmental Disorders, 36(1), pp. 27–43. doi: 10.1007/s10803-005-0040-7.

Okabe, M. and Ito, K. (2008) Color Universal Design (CUD): How to make figures and presentations that are friendly to colorblind people. Revised 24 September 2008. Available at: https://jfly.uni-koeln.de/color/.

Pardue, L.S. and McPherson, A. (2019) 'Real-time aural and visual feedback for improving violin intonation', Frontiers in Psychology, 10, article 627. doi: 10.3389/fpsyg.2019.00627.

Salmoni, A.W., Schmidt, R.A. and Walter, C.B. (1984) 'Knowledge of results and motor learning: A review and critical reappraisal', Psychological Bulletin, 95(3), pp. 355–386.

Schmidt, R.A. and Lee, T.D. (2011) Motor control and learning: A behavioral emphasis. 5th edn. Champaign, IL: Human Kinetics.

W3C (2023) Web Content Accessibility Guidelines (WCAG) 2.2. W3C Recommendation, 5 October. Success Criterion 1.4.1, Use of Color. Available at: https://www.w3.org/TR/WCAG22/.

Winstein, C.J. and Schmidt, R.A. (1990) 'Reduced frequency of knowledge of results enhances motor skill learning', Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(4), pp. 677–691. doi: 10.1037/0278-7393.16.4.677.