



OPEN ACCESS

EDITED BY

Anna Xambó,
Queen Mary University of London,
United Kingdom

REVIEWED BY

Charalampos Saitis,
Queen Mary University of London,
United Kingdom
Gerard Roma,
University of West London, United Kingdom

*CORRESPONDENCE

Benjamin Keith Bacon
✉ bacon@campus.tu-berlin.de

[†]These authors have contributed equally to this work

RECEIVED 21 March 2025

ACCEPTED 04 September 2025

PUBLISHED 21 October 2025

CITATION

Bacon BK, Auri L and McKee HA (2025)
Approaches to notation for embodied
engagement with a novel neural
network-based musical instrument.
Front. Comput. Sci. 7:1597806.
doi: 10.3389/fcomp.2025.1597806

COPYRIGHT

© 2025 Bacon, Auri and McKee. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Approaches to notation for embodied engagement with a novel neural network-based musical instrument

Benjamin Keith Bacon^{1*†}, Leonardo Auri^{1†} and
Holly Ambrozic McKee²

¹Audio Communication Group, Technical University of Berlin, Berlin, Germany, ²Digital Health, Hasso Plattner Institute, University of Potsdam, Potsdam, Germany

Introduction: The growing complexity and breadth of sonic possibilities enabled by sound synthesis technologies give rise to significant challenges for the notation of music, especially in light of emerging neural network-based paradigms. Prescriptive vs. descriptive notation has emerged as a paradigm with relevance to this challenge.

Methods: Experienced musicians ($n = 11$) were asked to compose for a novel neural network-based digital musical instrument and were prompted to produce descriptive and prescriptive graphic notations.

Results: Group differences in task conceptualisation were observed, while further analysis revealed perpendicular dimensions of variation in the resulting musical notation, which were associated with perceived creativity support.

Discussion: Based on this analysis, a conceptual framework is proposed that suggests useful strategies for music composition and creative endeavours.

Summary: Beyond the representation of sound and/or action, use of abstraction and metaphor emerge as strategies for music notation, with consequences for support of creative work.

KEYWORDS

graphic music notation, prescriptive and descriptive notation, digital musical instruments, neural-network synthesis, embodied music cognition, creativity support, notation design strategies, human-computer interaction

1 Introduction

Music notation is defined as “a visual analogue of musical sound, either as a record of sound heard or imagined, or as a set of visual instructions for performers” (Bent et al., 2001). Prescriptive vs. descriptive notation has emerged in the music notation literature as a paradigm for understanding and describing notational approaches, where descriptive notation intends to encode the “sound of a musical work,” whereas a prescriptive one “informs us of the method of producing the sound” (Seeger, 1958; Kanno, 2007). On a cognitive level, this distinction might reflect a contrast between a sonic (descriptive) and an embodied (prescriptive) approach to musical thought and task execution. In particular, Kanno (2007) suggests that prescriptive notation, which focuses on representing the actions or steps required to produce a sound, may be better suited to notating complex timbres (i.e. tone colours). This distinction is particularly relevant in the context of Digital Musical Instruments (DMIs), whose capacity to produce timbres beyond the physical constraints of acoustic instruments challenges traditional (i.e. historical) notation systems, which were not designed with such flexibility in mind (Wishart, 1996; Hope, 2017; Magnusson, 2019).

We set out to study the variation in notational approaches induced by prompts engineered to elicit prescriptive vs. descriptive notational styles in order to better understand the impact of novel DMIs on musical compositional processes.

We asked musicians to work with a DMI with a fixed control layout mapped to the high-dimensional latent space of a RAVE neural network synthesiser (Caillon and Esling, 2021). Since no established playing technique or repertoire for this instrument exists, all performers meet it as a novel instrument regardless of their prior experience.

Although we observed per-group differences in the thought processes and conceptualisations with which participants approached the task, our analysis revealed additional dimensions of variation within participants' scores with significant repercussions for understanding and undertaking compositional processes.

1.1 Background

1.1.1 Notation for digital musical instruments

The existing literature suggests that notating for DMIs is especially challenging, given the plasticity of their interaction mappings and their tremendous flexibility in sound output (Magnusson, 2019; Frame, 2023). Unlike traditional acoustic instruments, whose long-term standardisation fostered broad agreement on tuning systems, scales, and the quantisation of pitch and timing (Bent, 2025; Gould, 2016), DMIs transcend the physical constraints that limit sound production. Consequently, traditional Western notation, rooted in a culture of fixed pitch and discrete rhythmic values, struggles to represent the fluid sonic behaviour enabled by contemporary technologies.

This challenge becomes particularly evident when considering the complexity of *timbre*, which may be understood as a “complex auditory attribute, or as a set of attributes, of a perceptually fused sound event in addition to those of pitch, loudness, perceived duration, and spatial position (Siedenburg et al., 2019).” Traditional notation evolved over time for fixed-timbre, acoustic instruments whose sonic palette is delimited by materials, resonant behaviour, and specific playing techniques. In contrast, DMIs can change timbre dynamically across the audible spectrum, yielding subtleties that are difficult to capture on a conventional stave (Frame, 2023).

Music notation has therefore traditionally had to represent instruments with relatively stable timbre (Seeger, 1958), so devoting its information bandwidth primarily to the communication of varying pitch and rhythmic placement. As such, the proliferation of DMIs, along with their richer set of possibilities for manipulating timbre, brings existing inadequacies of traditional music notation for representing timbral variation into focus. Consequently, contemporary artistic practice, as well as academic music-notation research, continues to explore alternative options for music notation, including approaches which more directly represent the method of actuation (see Section 1.1.1.2).

As composers navigate these important factors, contemporary notation systems continue to evolve, particularly in electronic and experimental music. They must balance the need for precise technical instruction with the representation of sonic outcomes, adapting to the expanding possibilities of performance and

composition. Understanding how this creative balance is achieved can provide us with key insights as to how the future of written music can adapt to an instrumental landscape increasingly impacted by novel digital technology.

This evolving dynamic sets the stage for our subsequent examination of analytical approaches to music notation and, afterwards, a detailed exploration of the descriptive-prescriptive paradigm.

1.1.1.1 Paradigms of music notation analysis

Scholarship on contemporary music notation takes approaches falling into several related camps emerging from distinct historical traditions. Here, we focus on theories that have developed since the proliferation of electronic musical instruments in the post-war period. Semiotic models (Nattiez, 1990) examine how notation encodes information, analysing icons, indices, and symbols to uncover the visual codes performers interpret. Performance studies (e.g. Cook, 2013) approach scores as rehearsal scripts, where gestures, movement, and bodily negotiation fill in meanings only partially specified on the page. The experimental and indeterminacy tradition (as seen in the works of Cage and Jandl, 1961) treats notation as a catalyst for open form and improvisation, employing chance procedures and inviting interpretation through minimal or unconventional instructions. Philosophical and ontological work, especially that of Goodman (1976) questions what constitutes a musical “work,” contrasting notated objects with evolving processes of performance and realisation.

Building on these traditions, the descriptive-prescriptive paradigm articulated by Kanno (2007) offers a practical and still widely used framework for analysing how notation operates in performance (e.g. Yamaguchi et al., 2022; Lavastre and Wanderley, 2021; Bacon, 2022; McKay, 2021). It makes a functional distinction between notations that represent the intended sounds and those that represent the intended actions of a given composition. Among the existing models, Kanno's framework aligns most directly with embodied performance research, as it foregrounds the performer's body as the mediating link between visual mark and musical realisation. For this reason, we adopt Kanno's continuum as the primary analytic lens in this study and expand on its application in relation to embodied music cognition in Sections 1.1.1.2 and 1.1.2.

1.1.1.2 Prescriptive and descriptive notations

The prescriptive and descriptive notation paradigm was first introduced by Charles Seeger in an ethnomusicology context (Seeger, 1958). Seeger defined descriptive notation as a system that represents the resultant sound of a musical work, and prescriptive notation as one which specifies the actions required to produce that sound. He applied this framework successfully in analysing how oral traditions interact with written notation, highlighting the different ways in which musical information is communicated and interpreted.

Building on Seeger's work, Kanno (2007) refined this distinction within the context of modern and experimental music, where unconventional techniques and electronically generated sound regularly challenge traditional notational conventions. In such settings, descriptive notation allows performers to infer sonic outcomes, while prescriptive notation provides explicit

performance instructions for sound production. As contemporary composition increasingly incorporates complex timbral, gestural, and electro-acoustic elements, this distinction becomes particularly relevant to our study.

To clarify this distinction in terms of our study:

- Descriptive notation: represents what the music should sound like, informing performers of the intended sonic outcome.
- Prescriptive notation: specifies how the sound should be produced, providing explicit instructions on performance actions or techniques.

This functional differentiation applies to both traditional and graphic notation systems. In traditional staff notation (i.e., Western notation), a Bach violin partita serves as an example of descriptive notation: the score encodes pitches and rhythms with such specificity that a trained musician can accurately imagine the music without hearing it played.

Similarly, graphic notation can employ both prescriptive and descriptive characteristics. A score that utilises abstract shapes or textural symbols to indicate sonic qualities operates as descriptive notation, as it conveys general sound characteristics without prescribing specific instrumental actions. Conversely, a graphic score that illustrates hand movements, sensor interactions, or step-by-step performance gestures functions as prescriptive notation, guiding the performer through execution rather than merely describing the intended sound.

The differentiation between prescriptive and descriptive notation made in our study should also be understood in terms of music notation's function as a communicator of musical ideas to humans, and not as representations aimed either at communicating acoustic (non-musical) information or at machine reproduction. In this sense, a representation such as a spectrogram, while practical for acoustic analysis or machine reproduction, cannot be understood as music notation in the sense of this body of literature, as it does not provide a form interpretable for human audition (Gerhardstein, 2002) or actuation.

1.1.2 Embodied sound and gestural interaction

The field of embodied music cognition addresses the interplay between sonic and somatic processes in musical cognition, and as such constitutes an important link between music cognition and research on prescriptive and descriptive notation. Consideration of the functions of various cognitive modalities in musical cognition informs the question as to which notational approaches might most effectively address musical needs. Embodied music cognition understands bodily movement and gesture to be intrinsic to musical activity and the perception of musical phenomena (Maes et al., 2014). For example, motor areas of the brain show activation when individuals are merely asked to imagine listening to music without any actual sound present. In this case physical gestures are shown to be essential to the expression of music, serving as supporting mechanisms which help structure performances (Godoy and Leman, 2010).

Consideration of the multi-modal nature of musical experience and embodied music cognition has important consequences for both music pedagogy and performance. Embodied music cognition research has shifted attention from solely perceptual foci to

interactive settings, indicating the complex processes of emergent pattern building, gestural enactment, and human expressiveness which shape our experiences with music (Leman et al., 2018). Our study seeks insights into notational approaches with reference to embodied musical cognition processes.

1.1.3 Digital musical instruments

A distinguishing feature of DMIs is the arbitrary nature of their interaction mappings (Hunt et al., 2002). Thus, the success of a DMI is dependent upon its mapping scheme, which both grounds its interactivity and defines the expected outcomes and sounds of the instrument itself. Key elements of a successful instrument (i.e. one which supports human creativity) are immediacy, accessibility to beginners, and potential for long-term development of proficiency. New users should be empowered to create and manipulate sound right away (cf. Wessel and Wright, 2002, a “Low ‘Entry Fee’ with No Ceiling on Virtuosity”). By this token, instruments must also allow players to grow and continue to develop their skills, allowing an uncapped skill development of potential expert players.

The exploratory nature of DMI research gives rise to both opportunities and challenges. This openness positions DMIs as creative tools without predefined outcomes, yet they must still scaffold interaction so that a relationship between instrument and player can emerge. Earlier commentaries framed this requirement through the lens of *third-wave* Human-Computer Interaction (HCI), which emphasises designing technology to support meaning-making and rich experience (Tanaka, 2019). More recent scholarship in *fourth-wave* or *entanglement* HCI, however, argues that designers, artefacts, and users are mutually constituted within socio-technological relationships (McPherson et al., 2024). Acknowledging these complexly intertwined relationships extends the analytical frame beyond isolated user-instrument dichotomy and aligns DMI design with contemporary debates in HCI.

Processes for developing DMIs have come from multiple approaches, with some advocating for leveraging expert techniques already present in the population of potential players (Morreale and McPherson, 2017; Cook, 2009). For example, designers try to access gestural knowledge and physical skills that can be tapped into by borrowing the design of an existing instrument (Gurevich and Muehlen, 2001). Other instruments have been designed to provide a more ambiguous approach through simple physical forms paired with complex sensor mapping, allowing the player to construct their own meaning through interactions with the instrument (Hunt and Kirk, 2000).

A substantial body of DMI research demonstrates that deep, practice-led insights often emerge from studies centred on a single instrument or a single expert performer. Classic examples range from Waisvisz's pioneering Hands gloves (Waisvisz, 1985) and Mathews' Radio Baton (Mathews, 1991), through the widely influential reacTable tabletop synthesiser (Jordà et al., 2007) and Palacio-Quintin's longitudinal Hyper-Flute auto-ethnography (Palacio-Quintin, 2008), to focused redesign work such as Cook's SqueezeVox Maggie (Cook, 2009). Magnusson (2011) explored prescriptive and descriptive strategies in the context of a single live-coding performance interface. More recent practice-based investigations, such as McPherson and Kim's (2012) exploration of the “second-performer problem” with the Magnetic Resonator

Piano and Eldridge and Kiefer's (2017) self-resonating feedback cello, show that small-n, instrument-centred studies remain a recognised and productive methodology. Collectively, this lineage establishes that rigorous qualitative analysis of a single performer-instrument ecology can surface design principles, reveal performer-technology co-adaptations, and inspire subsequent multi-user research, thereby validating the methodological stance adopted in the present work.

1.1.3.1 Interactions with neural network synthesis

The integration of NN-based sound synthesis and other Machine Learning (ML) techniques in DMIs has opened up new avenues for expressive musical performance, especially in terms of timbre generation. Previous work has shown that NN synthesis models can produce realistic models of acoustic instruments while also allowing for real-time timbre transfer between sound sources (Engel et al., 2017; Ackva and Schulz, 2024). Beyond synthesis, these models have also been explored for expressive control: for instance, Zheng et al. (2024) demonstrated a mapping strategy that links visual art to the latent spaces of audio synthesis models.

Despite these creative possibilities, ML-based synthesis techniques present several challenges, including issues of user control, transparency, and predictability. The so-called “black-box effect” limits users’ ability to fully understand or anticipate system behaviour, while the complexity of parameter mapping can make fine-tuned control difficult (Briot et al., 2020). These challenges highlight the need for further research into how musicians engage with neural synthesis models in real-world performance and compositional settings.

1.1.3.2 Environmental Instruments

In our study, we employ a novel musical instrument called Environmental Instruments introduced by Auri and von Coler (2025). This instrument integrates gestural control with NN-based sound synthesis, combining an intangible interface reminiscent of the theremin with a fixed, deterministic, non-linear mapping scheme. This design allows us to systematically compare gesture-centric behaviour and embodied musical interactions across a range of participants.

By employing a novel instrument, one which less invites preconceptions of its use and sound (“blank slate”), we aim to gain insight into our research questions disentangled from existing notational practices and instrumental associations. Through a lack of assumptions around use and sound, new notational approaches are forced in participant responses.

In the original study, musicians reported both frustration and enjoyment when discovering the instrument's behaviour, often viewing its unpredictable mappings as a source of artistic possibility. They also explored a broad sonic palette, from natural timbres to synthetic, “unnatural” textures. Building on these insights, we examine how participants navigate this NN-based instrument through prescriptive and descriptive notational tasks in order to explore the cognitive strategies they employ.

We build upon two key conclusions from the original work: first, that musicians should be given extended practice to master novel Artificial Intelligence (AI)-driven instruments, given their unique challenges and the unprecedented avenues for expression. Simplifying a complex control system would diminish insights into

possibilities opened up by this novel technology. Second, the study underscores the value of exchange between performers. Sharing new techniques, concepts, and approaches, especially in group settings, may accelerate both individual skill development and the “collective evolution of AI-driven musical practices.” To foster this, our research incorporates a focus group setting, providing a space where musicians can exchange insights and push the boundaries of this emerging class of instruments together.

1.1.4 Situating our work between disciplines

Any inquiry into how novel DMIs intersect with compositional approaches (particularly through notational representation) inevitably draws on multiple fields. Like modern DMIs, graphic notation challenges conventional notions of correctness, as both lack predefined “right” or “wrong” ways of being interpreted or played. This inherent openness invites a broader exploration of notation as both a representational system and a tool for creative expression.

The interdisciplinary nature of such research is reflected in venues such as the TENOR conference, founded in 2015 to examine technology's role in notation and representation. Because of its interpretive flexibility, graphic notation can be analysed in terms of its mappings between visual and musical elements, drawing parallels to research on mapping in DMIs (Bacon, 2022). Additionally, studies suggest that graphic notation, as a form of visual stimulus, can influence a performer's ability to learn a novel DMI (Wu et al., 2022).

Building on the distinction between prescriptive and descriptive approaches to notation, we align our perspective with third-wave HCI research, which emphasises human-centred experiences of technology in creative domains (Blackler et al., 2018). In addition, the mixed methods approach employed in our study offers unique insights into our research questions which tackle multiple complex subject areas (Migiro and Magangi, 2011). By examining how digital tools transform traditional approaches to notation and composition, we seek to understand how technology drives the evolution of new forms of music-making.

1.2 Research questions

To frame the investigation, this study is guided by three core research questions that examine the relationship between notation, creative engagement, and user interaction with a novel DMI:

- **RQ1:** How do prescriptive and descriptive notation methodologies shape user interaction, embodiment, and creative processes?
- **RQ2:** How does the use of prescriptive and descriptive approaches affect the design and complexity of the resulting graphic scores?
- **RQ3:** How does the absence of instrumental expectations influence users' experiences, perceptions, and gestures when working with a novel DMI?

These research questions form the foundation of our investigation into human creativity at the intersection of novel

technological tools such as neural network synthesis and traditional artistic practices like musical scoring. By exploring this dynamic relationship, we aim to provide insights into the evolving landscape of human-centric creativity as technology and tradition continue to shape musical experiences.

1.3 Contribution

The processes behind the creation and implementation of prescriptive and descriptive notation, particularly in the context of NN-based instruments and electronic instruments more broadly, remain unknown. Historically, the descriptive and prescriptive definitions of notation have been predominantly applied within the domain of music theory. This paper aims to extend these definitions into the field of HCI, exploring their relevance and robustness when applied in user studies and focus group settings. By examining these fundamental definitions in practice, we aim to evaluate how well they hold up in contexts involving the creation of notation for novel DMIs.

This study highlights the unique challenges and opportunities presented when designing notations for instruments that prioritise complex timbral changes. Furthermore, it demonstrates that these fundamental definitions provide a distinctive departure point for understanding how individuals conceptualise and create new notations in environments that lack predefined expectations or conventions. By bridging music theory and HCI research, this work sheds light on how people approach the process of creating notations for new and innovative musical systems.

Finally, we position this work as generative, theory-building research (Eisenhardt, 1989). We set out to inductively surface concepts that can frame future inquiry. By engaging hands-on with a still-novel NN-based DMI, the study enriches the communal repertoire of experiences and practical know-how required for this emerging design space, an outcome that could not be achieved through abstract comparison alone. In doing so we extend a line of open questions identified in earlier exploratory work with the same instrument (see Section 1.1.3.2), transforming them into empirically grounded propositions that subsequent large-scale or comparative studies can refine. Thus, the present research contributes not only its immediate findings on prescriptive and descriptive notation, but also a scaffolding of shared practice on which the NN-instrument community can continue to build.

2 Experimental setup

In our study, practising musicians with diverse musical and cultural backgrounds were recruited and issued with our Neural Network (NN)-based instrument, introduced above, to take home and experiment with in their own time. Participants were asked to compose a short *étude* for the instrument. The prompts given were designed to elicit either a prescriptive or a descriptive approach to the notation, depending on their experimental group.

The participants were then brought together in a focus group setting to discuss their creative process and the challenges and possibilities they faced in developing a notation system for the novel NN-based instrument. Finally, we employed three complementary analytical approaches to understand how participants engaged with

the instrument and task in terms of both their personal experiences and compositional strategies, and to compare the resulting musical scores in terms of their mappings and visual features.

2.1 The instrument

We build on the NN-based Environmental Instruments introduced by Auri and von Coler (2025). This instrument employs the RAVE architecture (Caillon and Esling, 2021), a temporal convolutional neural network (Bai et al., 2018) designed for real-time audio reconstruction tasks (Figure 1a). The RAVE architecture was trained on a birdsong dataset (Cornell Lab of Ornithology, 2020), after which movement input was interfaced with the decoder, allowing players to explore audio reconstruction of the network's latent space in real time (Figure 1b). The authors elected to freeze the instrument in its earliest playable state, in order to “avoid biases and gain insights” specifically into the nature of NN-based musical instruments (Auri and von Coler, 2025).

For the present study, the instrument was re-implemented as a JUCE standalone application.¹ Similarly to the aforementioned work, *Nintendo Switch* controllers are utilised as a gestural interface. Our version preserves the original mapping and behaviour, enabling participants to install and use the instrument on their own computers while maintaining consistency with the original instrument.

In the instrument's current form, the 128-dimensional latent space of the RAVE decoder was reduced to a 12-dimensional space via rank reduction and principal component analysis (see Caillon and Esling, 2021, Section 3.2). Scaled raw accelerometer (3-dimensional) and gyroscope (3-dimensional) signals from the left and right controllers were mapped directly to the 12-dimensional latent space reduction of the RAVE decoder.

In this mapping, different gestures yield a broad range of sonic outcomes, spanning near silence to recognisable or highly abstract bird-like timbres. Vigorous shaking often triggers variable or extended silent passages, while subtle movements can produce nuanced shifts within specific “regions” of the latent space. These regions may correspond to identifiable sounds from the training corpus, significantly altered mutations of those sounds, or entirely unrecognisable sonic textures.

Thus, we present participants with a novel DMI that provides a space for imagination and exploration. As Hunt et al. (2002) suggest, exploring prototype instruments opens a space for performers to experiment with different musical pathways, fostering creative engagement. This aligns with Buxton (2007)'s perspective that early-stage prototypes allow room for discovery and creativity, rather than imposing rigid constraints.

2.2 Participants

A total of 11 participants were recruited through an open call to email list servers from the Catalyst Institute for Creative Arts and Technology, as well as the Technical University of Berlin, and

¹ See <https://github.com/leoauri/environmental-instruments-published> for the standalone implementation's code.

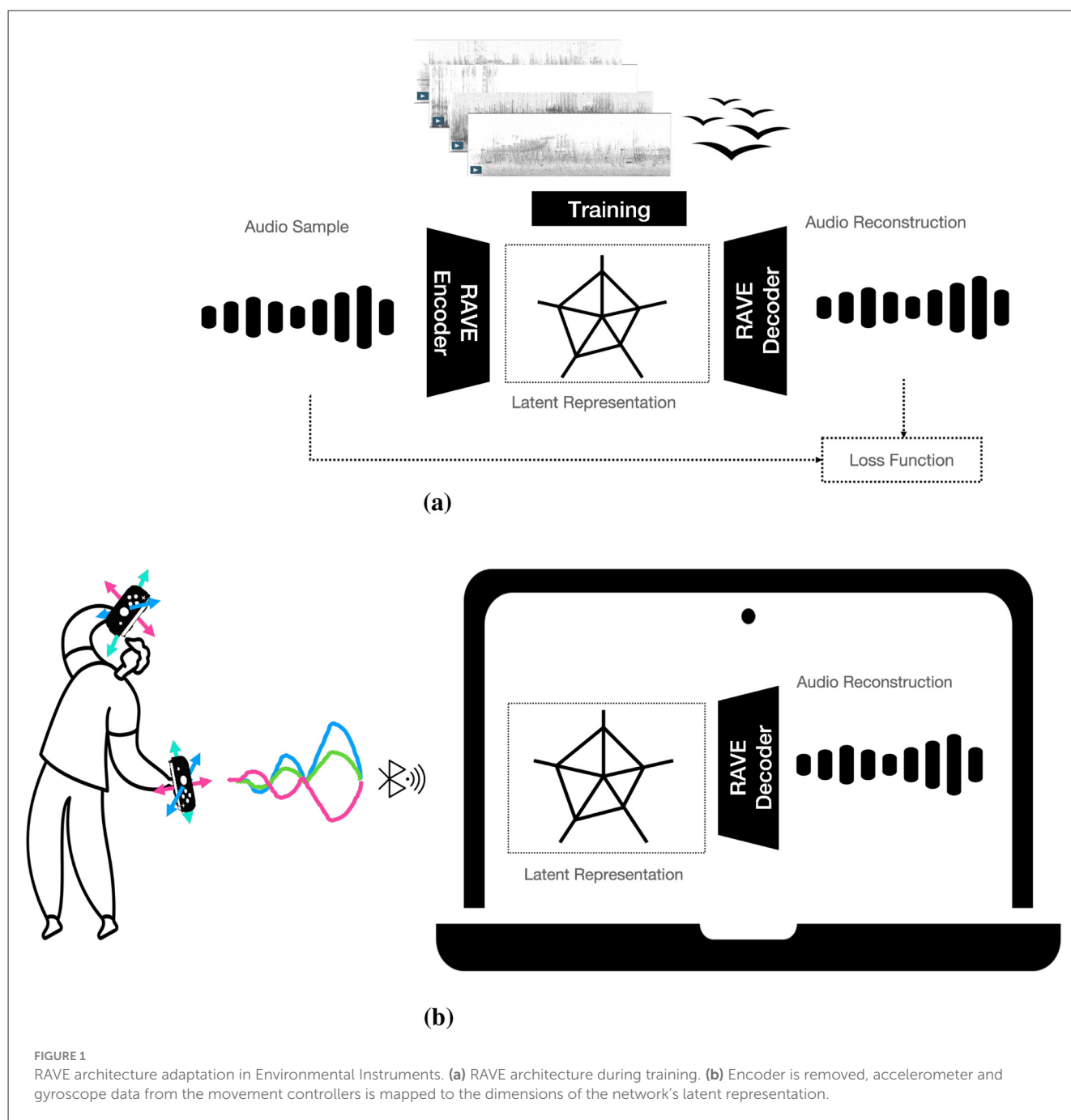


FIGURE 1

RAVE architecture adaptation in Environmental Instruments. **(a)** RAVE architecture during training. **(b)** Encoder is removed, accelerometer and gyroscope data from the movement controllers is mapped to the dimensions of the network's latent representation.

participated on a voluntary basis. We included actively practising amateur and professional musicians with a variety of backgrounds and specialisations.

Two of the participants who responded had participated in a pre-study with the instrument (Auri and von Coler, 2025). This study gave participants one approximately half-hour session using the instrument. Since this still allowed all participants comparable experience with the instrument before task completion, it was determined to admit these participants for the current study. Finally, none of the participants reported experience with Nintendo Switch controllers outside of the context of this instrument.

We wished to match our experimental groups in terms of musical abilities as closely as possible. To this end, participants were given a Goldsmiths Music Sophistication Index (GMSI) assessment upon recruitment, which evaluates a 5-dimensional model of musical sophistication by “explicitly considering a wide range of facets of musical expertise” (Müllensiefen et al., 2014). Participant results from the GMSI can be seen in Table 1. By treating these scores as a 5-dimensional Euclidean space, the blossom algorithm (a graph-theoretic method for finding a best-fit set of pairings by addressing problematic cycles that can arise, see Edmonds, 1965) was used to find a minimally distant pairing of participants. One

TABLE 1 Study participants.

Participant	Pair	Group	AE	PA	MT	SA	Em
P1	A	P	5.56	6.56	6.57	5.14	5.17
P2	E	P	4.78	5.78	6.14	4.86	5.67
P3	F	P	6.78	7.00	6.57	6.86	6.00
P4	C	P	4.22	5.00	4.71	4.14	4.83
P5	D	P	6.11	5.00	4.43	3.00	6.67
P6	B	P	5.11	7.00	5.71	6.00	5.67
P7	C	D	3.78	5.89	3.00	3.71	5.33
P8	F	D	6.78	7.00	6.86	6.00	6.50
P9	E	D	4.33	6.11	6.57	5.57	6.33
P10	B	D	5.11	5.00	4.14	6.00	6.00
P11	A	D	5.89	6.11	6.29	4.00	6.17

Groups: P, prescriptive; D, descriptive. Dimensions of GMSI: AE, active engagement; PA, perceptual abilities; MT, musical training; SA, singing abilities; Em, emotions.

of each pair of participants was then placed into each of our two experimental groups.

Due to withdrawal from the study, one pair was left incomplete. However, mean GMSI “general sophistication” remained very close between groups ($D = 5.54$, $P = 5.57$).

This study received ethics approval from the “Kommission für Forschungsethik der Fakultät I” at Technical University of Berlin, ensuring compliance with ethical research standards. As with all human subjects research, participants retained the right to withdraw at any time, and their identities have been made anonymous to protect confidentiality.

2.3 Task

Participants received a copy of the instrument for their own use at least 1 week prior to their scheduled focus group meeting, with all participants receiving access within the same 1-week window. They were given the freedom to engage with the instrument as much as they wished in preparing their piece and completing the task. Each participant was assigned one of two prompts intended to guide their approach to creating a graphic notation. The prompts were designed to elicit either a prescriptive or a descriptive notation approach.

Participants in the prescriptive group received the following prompt:

Your graphic notation should aim to communicate the methods or actions required to perform your composition. In other words, your graphic notation should act as visual instructions or diagrams, detailing the gestures and movements necessary to realise the performance.

Participants in the descriptive group were given the instruction:

Your graphic notation should aim to communicate the intended sounds of your composition if it were performed. In other words, your graphic notation should visually represent the necessary information to imagine what it may be like to hear the piece itself.

In addition to creating a graphic score, participants were instructed to keep a journal documenting their engagement with the instrument, including their “observations, thoughts, and creative processes” (RQ1). No specific length requirements were imposed for these journal entries, allowing participants the flexibility to reflect on their experiences in a manner that suited their individual workflows.

To ensure that participants avoided conventional Western notation, the term graphic notation was used in the prompt text. Please replace with: Participants were asked to avoid colour in order to reduce visual complexity of scores, to avoid perceptual confounds of colour and sound (Ramachandran et al., 2020), and because we could not ensure equal access to coloured drawing/design materials across individuals. No participants in the focus group or elsewhere referenced the lack of colour as an issue or a hindrance to their creative process.

2.3.1 Creativity Support Index

Upon instrument collection and before beginning the task itself, participants completed the Creativity Support Index (CSI) paired-factor comparison sheet to assess their notation systems (Cherry and Latulipe, 2014). The CSI’s pre-task assessment asks respondents to compare the six major creativity support factors (exploration, expressiveness, immersion, enjoyment, results worth effort, and collaboration) in pairs, ranking their relative importance. These rankings establish a baseline weighting of their expectations and preferences regarding creativity support.

On the day of the focus group, before the discussion had started, participants completed a final CSI questionnaire, evaluating their experience based on the graphic notation they had created for the instrument. This post-task assessment aimed to capture participants’ sense of creativity support resulting from the notation process.

The questionnaire ranks each of the six major creativity support factors on a scale from 0 to 10, where 0 indicates no support, 5 is neutral, and 10 represents absolute support. Final CSI scores are calculated by weighting the ratings by participants’ pre-task paired-factor rankings, ensuring that the final results reflect both their experience and their individual creative priorities. This process produces a CSI score between 0 and 100, where a higher score indicates a greater perceived level of creativity support in relation to the tools and methods under review.

2.4 Focus group

We employed a focus group methodology (Breen, 2006) in order to encourage both discussion and a deeper exchange between players of the instrument. We incorporated time for participants to perform their scores and engage in an exchange around their

creative processes. To maintain a small group size that would allow for deeper engagement (Greenbaum, 2000; Krueger and Casey, 2015), the 11 participants were split into two focus group sessions. Session 1 included seven participants (P1, P2, P3, P6, P7, P9, P11), while Session 2 included four participants (P4, P5, P8, P10). Session 2 occurred 5 days after Session 1, but all participants had concluded their task by the first focus group.

The focus group took the form of a guided discussion. It began with feedback on the instrument including its affordances, followed by each participant describing and performing their score, and concluding with questions on their approach to notating their étude.

3 Analysis

We organise our results around the three analytical frameworks used in the study. We apply *Thematic Analysis* (Section 3.1) to participant statements to identify differences and similarities in approach and experience between experimental groups. Section 3.2 analyses the *Notation Design Space* (NDS) to show how mapping density and visual-musical choices diverged across prompt conditions. Together these lenses build a coherent account of how prompt framing shapes notational practice. Section 3.3 integrates a *Grounded Visual Taxonomy* with creativity support ratings, revealing dimensions of score variation orthogonal to the prescriptive-descriptive divide.

3.1 Thematic analysis

3.1.1 Methodology

We conducted a thematic analysis of both the focus group transcripts and participants' journals in order to explore participants' subjective experiences with the instrument, their perception of its affordances and constraints, and their decision-making processes when constructing notation. This qualitative analysis helps contextualise the data-driven approaches by revealing how participants conceived of their interaction with the instrument, how they mapped sonic properties to gestures, and subsequently adapted their notation strategies in the absence of traditional musical parameters.

The recordings from both focus groups were transcribed using the *Descript* software (<https://www.descript.com/>). Quotes from the focus groups, as well as participant journals, were then labelled with anonymised participant ID, experimental group, and pair. The research team carefully reviewed the transcripts while simultaneously cleaning the data.

The team then conducted thematic analysis using *Atlas.ti* software (<https://atlasti.com>), following a reflexive thematic strategy (Braun and Clarke, 2019), incorporating both open and axial coding (Corbin and Strauss, 2015; Fereday and Muir-Cochrane, 2006) on the focus group sessions and the experience journals provided by the participants.

Initial codes were generated as a team using a line-by-line analysis for the first focus group transcript, after which the coding was done iteratively by the three team members. The research questions were used as guides for identifying relevant quotes.

Once coding was complete, codes were sorted collaboratively into overarching themes and sub-themes. To gain insight into the creative process, quotes were grouped into reflections on instrument exploration, strategies for creating notation and potential future uses of the instrument. Similar codes were merged, resulting in a final code-book including sub-codes, such as “Expression through movement” and organised into code groups such as “Experiences of Embodiment.”

Below we explore the difference in approaches taken by the prescriptive and descriptive task groups in line with the research questions, while highlighting overarching themes.

3.1.1.1 Validation of prompt-induced differences in task conceptualisation

In order to validate that the differing prompts induced differences in how participants conceptualised the notation task, we performed a quantitative analysis of the coded statements. Codes were grouped into the categories “Gesture and Notation” and “Sound and Notation,” which respectively relate gesture and sound to notation. Bayesian binomial mixed-effects models with random intercepts for participant and statement were trained to examine the effect of prompt manipulation on statement production for both themes.

3.1.2 Results

Bayesian binomial mixed-effects models with random intercepts for participant and statement indicated that the prompt manipulation had a significant effect on statement production for both themes. For the “Gesture and Notation” theme, the posterior mean for the prompt effect in the prescriptive group was 0.72 (SE = 0.19), with a 95% credible interval of (0.35, 1.09). This corresponds to an odds ratio of 2.06 (95% CI: 1.41–2.98), indicating that participants receiving the prescriptive prompt were approximately twice as likely to make statements concerning “Gesture and Notation” as those receiving the descriptive prompt.

Furthermore, for the “Sound and Notation” theme, the posterior mean for the prompt effect in the descriptive group was 2.21 (SE = 0.31), with a 95% credible interval of (1.60, 2.82). This translates into an odds ratio of 9.14 (95% CI: 4.96–16.84), meaning that participants in the descriptive group were approximately nine times more likely to make statements regarding “Sound and Notation” compared to those in the prescriptive group.

These findings confirm that the prompt manipulation effectively altered task conceptualisation, providing a robust basis for further analyses.

3.1.2.1 Embodied perception of instrument

When discussing the exploration of the instrument and what it afforded, themes related to embodiment emerged, with code occurrence slightly higher for the prescriptive group.

The prescriptive group had a higher occurrence of the code group “physicality of instrument,” most significantly supported by mentions of “body discipline” and “sense-making with body.” In terms of conceptualisation of gesture, they considered movement to be “exposing” or “driving” the sound, whereas the descriptive group tended to see expression as producing the sound and opted for actions and gestures that “fit the theme of the piece.”

Whether prompted to focus on sound or gesture, participants often reflected that the expressivity of the performed piece stemmed from the movement, often relating it to dance and choreography.

3.1.2.2 Strategies for creating a notational language

There was a wide diversity in approaches taken to creating a new notational language by participants, including building sound or gesture libraries, gesture maps, and instruction manuals.

In the descriptive group, P11 started by creating a library of sounds, ... “approaching the notation system from the properties of the sound of the movements.” [P11]

The strategy of borrowing from language notation and iconography was also employed. Inspired by Arabic calligraphy, P5 employed an approach in which “gesture follows the shape of the word.” [P5]. P6 based their notation on aircraft marshal signals, using recognisable gestures to structure their piece:

I decided to construct a piece with movements based on the movement of aircraft marshals. I could find icons for these movements, and after trying some combinations, I decided to follow/create a structure or sequence of signs rather than making the sonic reactions or sonic manifestation lead the way. [P6]

3.1.2.3 Metaphors for sound and gesture

Both groups frequently used naturalistic metaphors to describe their gestures and compositions, drawing on imagery such as “being a bird” [P4], moving like a wave [P8] or a meditation at the beach (represented by cricket and seagull sounds) [P7].

In the descriptive group, composition often began with sound exploration, with gestures following the imagery evoked by the sound:

I thought of the entire thing based on waves because that’s what the sound reminded me of ... I often used sweeping, wave-like gestures to reflect the contours I heard.” [P8]

For another participant, sound evoked strong emotions and vivid imagery:

The sounds felt scary, like creatures being unleashed underground... some were brutal-like the squeals of dying pigs-demanding a matching physicality. [P10]

They also reflected on how their movement then in turn shaped their sonic perceptions:

I was trying to capture how my movement’s emotionality influenced the images I heard. [P10]

In contrast, the prescriptive group usually placed movement at the forefront, using metaphors to shape their gestures. P4 designed a “visual poetry” of migrating birds, with gestures reflecting grace, lightness, swiftness, and fluidity in connection to the sounds.

3.1.3 Discussion

Prescriptive and descriptive groups appeared to differ in their conceptual approach to the creative task and process of exploring the possibilities of the instrument, with participants in the prescriptive group more driven by embodied processes, and descriptive tending to structure their creative process thematically (RQ1).

In Section 3.1.2.2 we can see how participants’ notational strategies influenced score design (RQ2), spanning across descriptive (e.g., P11 starting with sound properties), prescriptive (e.g., P6 using a predefined gesture system) and hybrid approaches (e.g., P5’s calligraphy-inspired gestures). These choices had a visible impact on the structure, interpretation, and complexity of their resulting graphic scores, which highlights how different starting points can shape both the visual language and the performer’s interactions (RQ1).

Section 3.1.2.3 shows us that with the descriptive group, the creative process began with sound, with participants drawing from naturalistic metaphors to guide their gestures (i.e. waves, underground creatures) (RQ3), forming a workflow from sound to movement (RQ1). In contrast the prescriptive group started with movement, using imagery (i.e. migrating birds) to shape their gestures (RQ1) which then connected to sound. Whether gestural or sonic, participants employed metaphor to guide their work in the absence of familiar instrumental or sonic expectations (RQ3).

3.2 Notation Design Space analysis

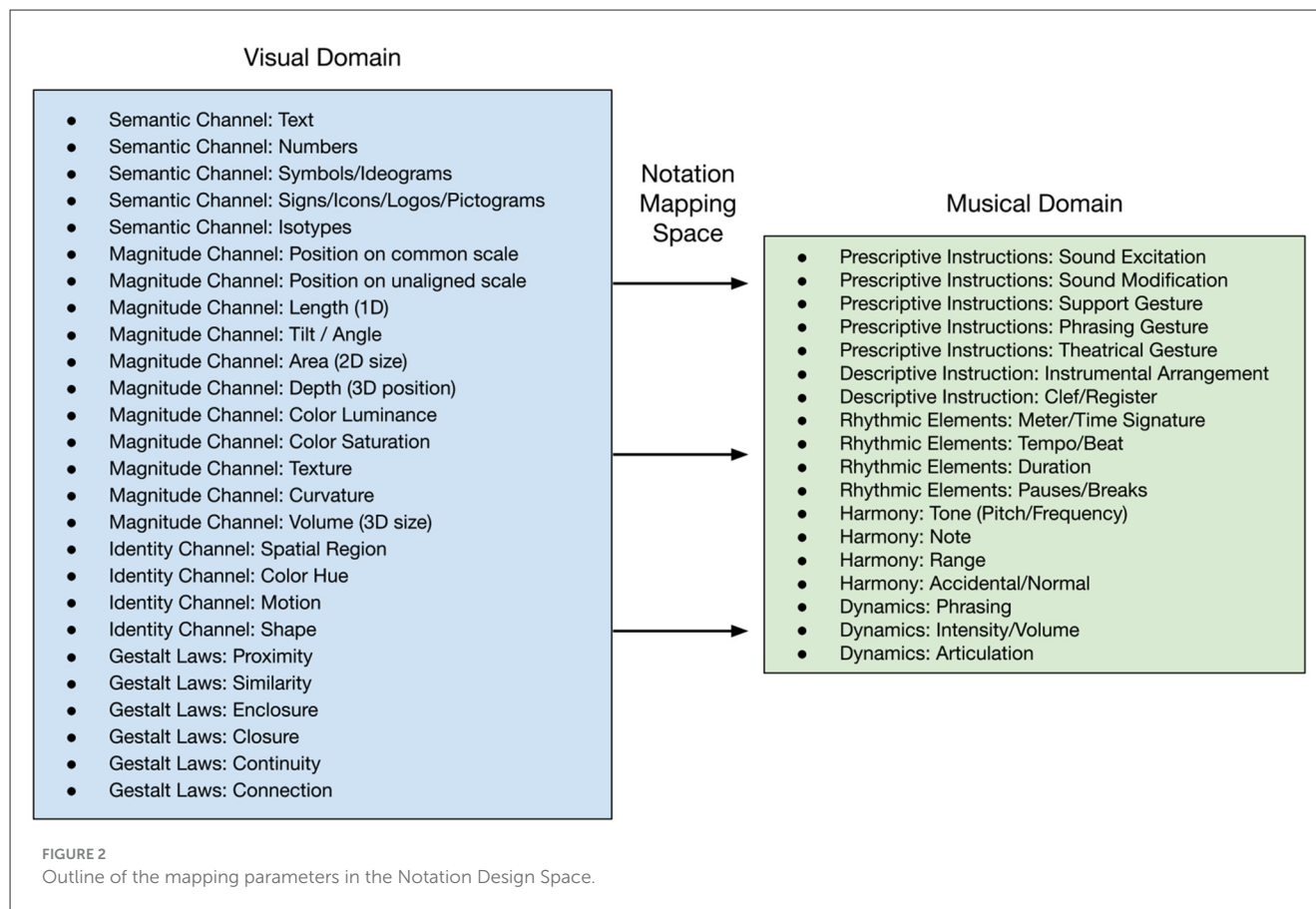
3.2.1 Methodology

Building on previous work that introduced the Notation Design Space (NDS) (Miller et al., 2018; Bacon, 2022), our analysis investigates the mappings between graphic channels and musical compositional features. Consistent with this approach, we traced unique one-to-one mappings between graphic channels and musical features, allowing for multiple mappings to a single feature. For example, if both luminosity and curvature were mapped to dynamics in a score, both mappings were recorded.

A comprehensive matrix was constructed to enumerate all possible graphic-to-musical connections, yielding 468 unique comparisons between 26 graphical features and 18 musical features. For each participant’s graphic score, the presence of a mapping was coded present or absent. We then computed the overall mapping counts per participant as well as group-level means and variances across musical and visual domains and their subdomains. Figure 2 outlines the mapping parameters in the Notation Design Space.

3.2.2 Results

The results of our NDS analysis show differences between the descriptive and prescriptive groups in terms of mapping density and variance (see Table 2), as well as a distinct use of the subcategories of the visual and musical domains (see Figures 3a, b). In our study, the prescriptive group exhibited a group mean of $M = 16.5$ mappings per participant (with a standard deviation



of $SD = 3.39$), while the descriptive group recorded a higher group mean of $M = 20.6$ mappings (with a standard deviation of $SD = 10.36$). These differences in mapping density indicate that participants in the descriptive group, on average, produced more mappings, although the variance was also higher.

In addition to the overall mapping counts, a breakdown by domain subcategories provides further insight into group differences in notation design choices. For musical features, the *Prescriptive Instructions* subcategory was used most frequently by both groups, with a mean of $M = 8.2$ mappings per participant in the descriptive group and $M = 7.83$ in the prescriptive group. Differences were found in the *Harmony* (D group $M = 2.40$, P group $M = 0.83$) subcategory, and the *Dynamics* subcategory also showed a slight difference (D group $M = 6.40$, P group $M = 4.00$). For visual features, the descriptive group predominantly used the *Semantic* channel (with a mean of $M = 14.6$), whereas the prescriptive group showed a slightly higher usage of the *Magnitude* channel (with a mean of $M = 6.5$).

Overall, the descriptive group exhibited a higher overall mapping count and greater variability. The use of the musical subcategories did not differ significantly between groups with the possible exception of the *Dynamics* and *Harmony* musical subcategories. However, we found interesting differences between groups in the visual domain, such as the high usage of the *Semantic* channel in the descriptive group and the relatively higher use of the *Magnitude* channel in the prescriptive group, suggesting underlying differences in mapping strategies.

3.2.3 Discussion

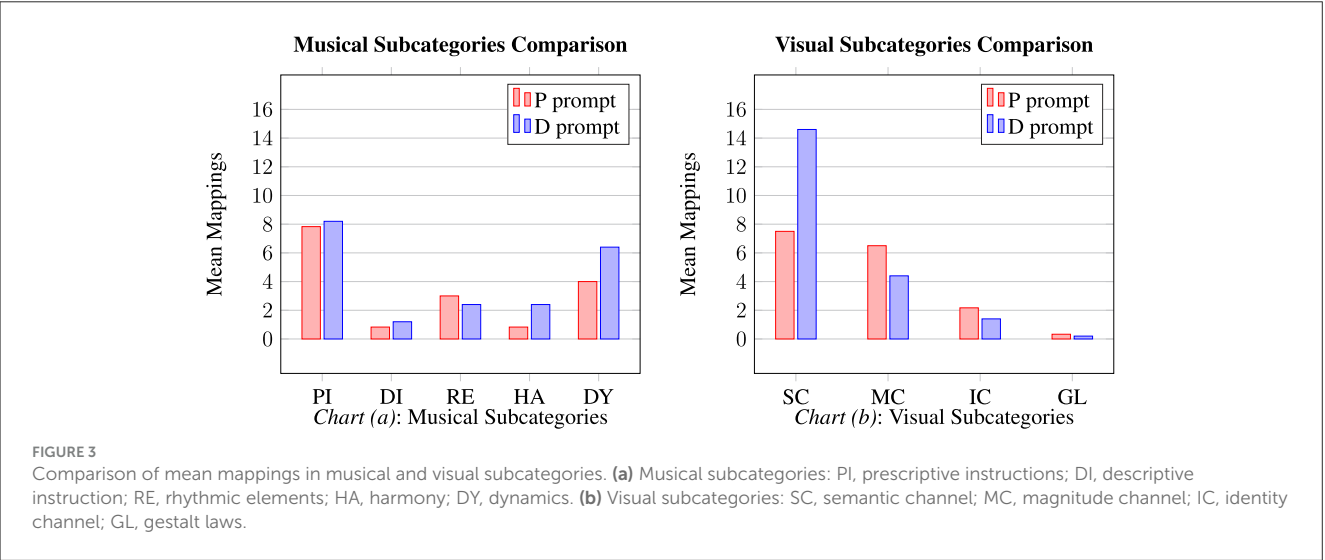
This analysis highlights important differences between the descriptive and prescriptive groups, particularly in terms of mapping density, variability, and visual strategies (RQ2). Participants given descriptive prompts produced notably more mappings on average than those receiving prescriptive prompts. This suggests that descriptive prompts encourage participants to generate a broader array of graphical-to-musical associations, potentially due to the interpretative flexibility and openness inherent in descriptive instructions. This is perhaps best exemplified by P10 (see Figure 4), whose score employed a rich diversity of graphical elements used to convey specific sound-image metaphors for performance interpretation.

Beyond the overall mapping counts between the two groups, examining the musical and visual subcategories reveals potential differences in design strategies employed by the prescriptive and descriptive groups. While musical mappings remain relatively balanced across subcategories, we observe interesting contrasts between the groups in the visual domain.

The descriptive group used an average of 14.6 mappings per participant in the *Semantic* channel, nearly twice as many as the 7.5 mappings per participant in the prescriptive group. This suggests that the descriptive prompt may have encouraged a more holistic approach to graphical representation, allowing for broader conceptual associations and greater interpretative flexibility. Conversely, in the *Magnitude* channel, the prescriptive group used 6.5 mappings per participant, compared to 4.4 in the

TABLE 2 Graphical to musical feature mapping counts in the graphic scores of each participant; group mean and standard deviation.

Prompt group	Prescriptive						Descriptive				
Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Feature totals (FT)	18	18	16	15	11	21	22	19	16	37	9
Group mean of FT	16.5						20.6				
Standard deviation of FT	3.39						10.36				



descriptive group, indicating a stronger emphasis on perceptual elements and analytical design, likely influenced by the need for clear instructional communication (RQ2).

While the small sample size limits firmer conclusions, these findings highlight promising directions for further research into information density and graphical display techniques across descriptive and prescriptive notation paradigms in digital musical instruments.

3.3 Grounded visual taxonomy and principal component analysis

3.3.1 Methodology

We developed a taxonomy of score characteristics and applied Principal Component Analysis (PCA) to reveal structural patterns between the graphic scores themselves. By systematically describing and encoding the visual features of each score, we aimed to extract key variables and analyse whether distinct notation paradigms (descriptive vs. prescriptive) emerged in a latent space.

We adopted a grounded, bottom-up approach (Corbin and Strauss, 2015; Charmaz, 2014) to categorise the notational elements in participants' scores, following precedents in notation research (Miller et al., 2018; Bacon, 2022). Two authors independently examined each score, noting the presence or absence of specific characteristics. To avoid trivial or overly narrow categories, we excluded any characteristic that appeared in all scores, in all but one score, or in only one score. The authors then compared and merged their lists, removing duplicates and confirming that

each characteristic captured a distinct element of notation. The resulting feature set and score categorisation is available.² This non-hierarchical taxonomy formed the basis of a PCA, designed to identify underlying dimensions of variation within participants' notational practices.

3.3.2 Results

We performed PCA of our taxonomy of notational elements in participants' scores. Figure 5 shows the cumulative variance explained by the components.

Examination of the loadings led us to conclude that the first two components summarise meaningful overarching tendencies, while the third component and beyond primarily serve to untangle noise in the dataset. See Appendix 1.1 for a summary of taxonomy elements which loaded strongly onto the first two principal components. Figure 6 shows participants' scores individual placement in the PC1/PC2 space.

Furthermore subtle correlations with prescriptive and descriptive prompts should be observed, with the centroids of the groups falling in the upper-right and lower-left quadrants of this space respectively (marked with "+" in Figure 6).

3.3.2.1 Creativity Support Index

We collected CSI scores from participants, which are depicted in Figure 7. The creators of the CSI instrument suggest drawing a parallel to educational grading systems in the interpretation

² See https://osf.io/c5fq9/?view_only=fad445f2d1f146db9d9d02cf8d7ed6c6.



FIGURE 4

P10's score employs a rich diversity of graphical elements, and demonstrates the abstraction of musical signals, so asking for greater interpretative input from the performer.

of CSI scores, with, for example, a score above 90 indicating “excellent support for creative work,” or scores below 50 indicating tools which “do not support creative work very well” (Cherry and Latulipe, 2014). CSI scores are also reflected in Figure 6 by dot size.

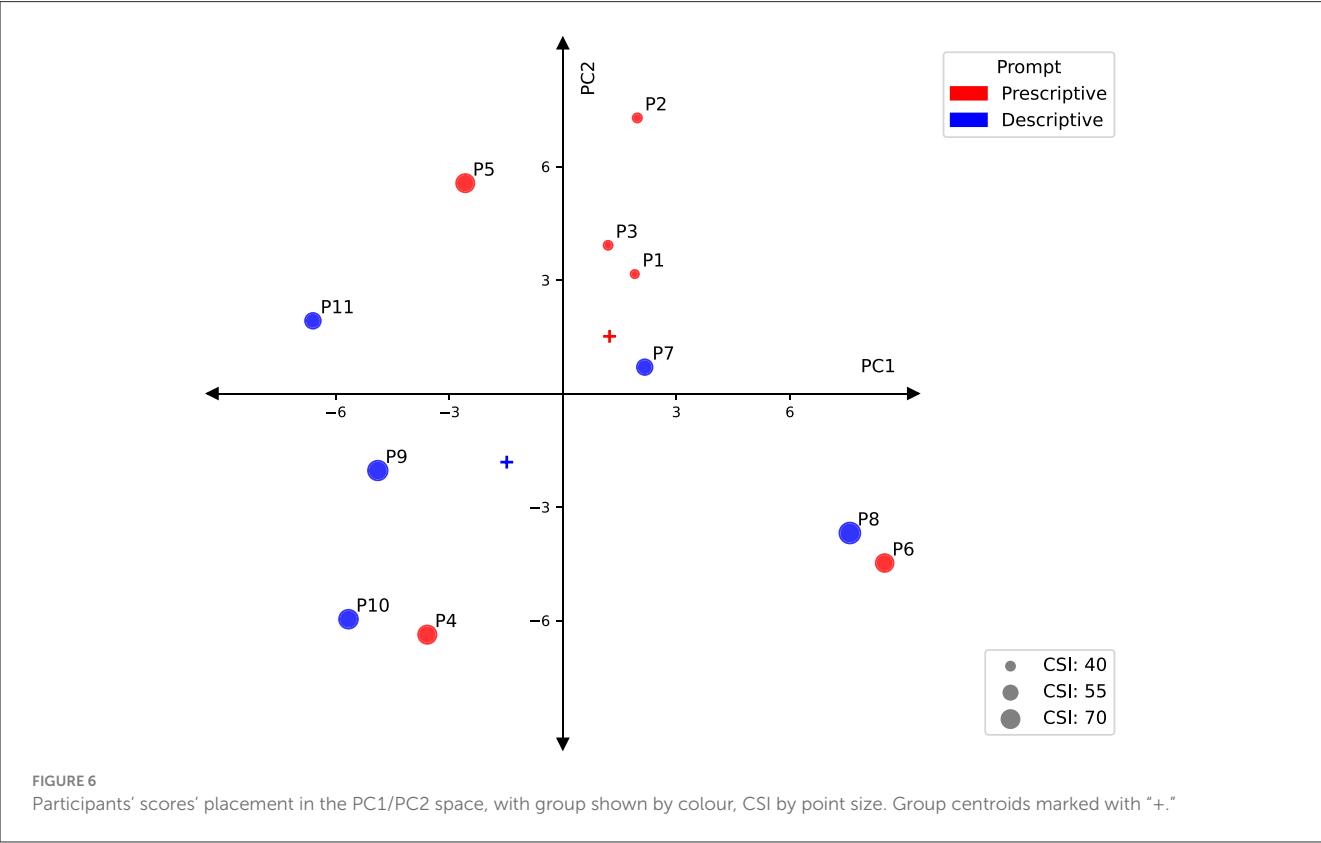
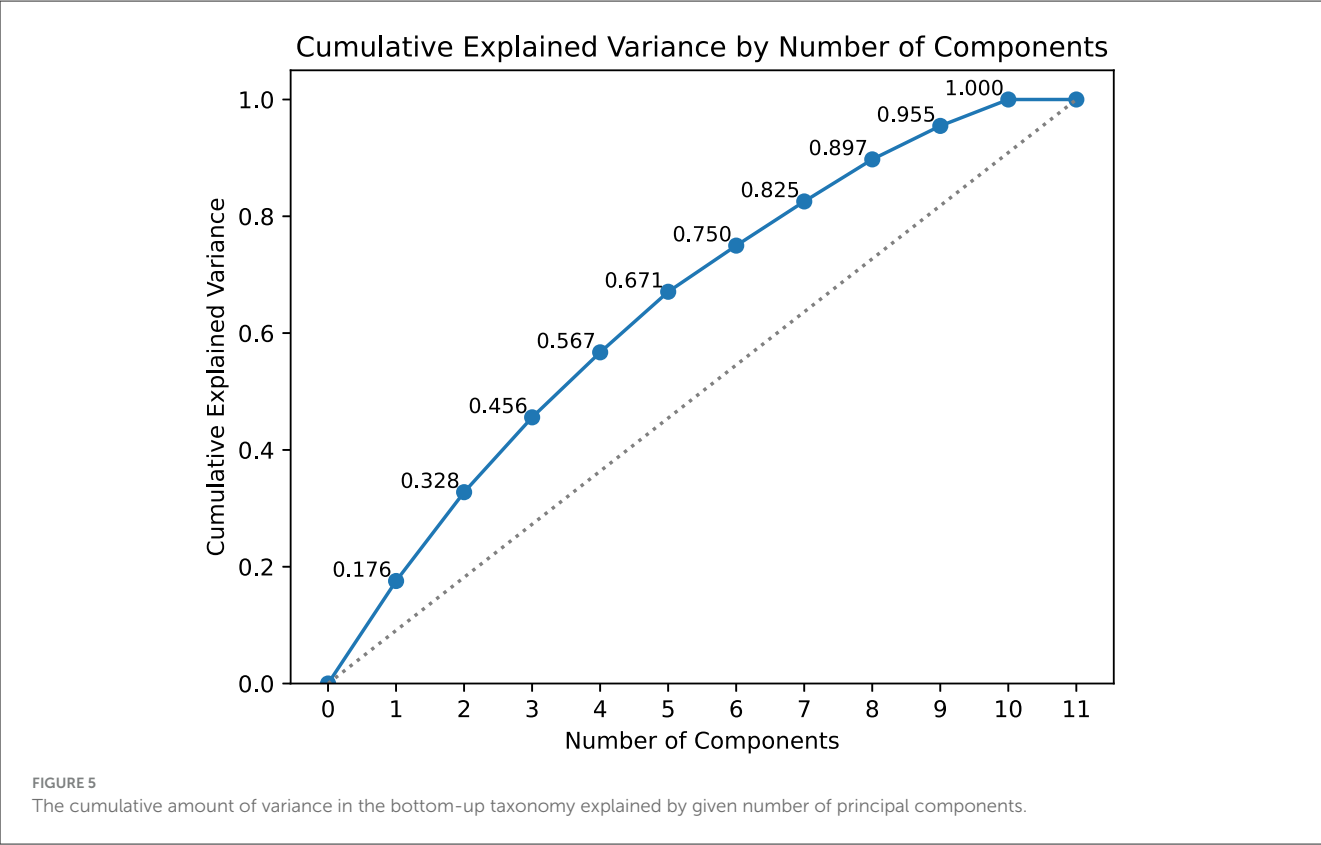
An ordinary least squares regression showed that the combined effect of quadrant membership (PC1 and PC2 both positive) significantly predicted CSI with $R^2 = 0.754$, $p < 0.001$ (see Figure 8). Correlation of CSI with either PC1 or PC2 individually was not statistically significant. See the discussion (Section 3.3.3.8) for further consideration of this result.

3.3.3 Discussion

The existing music notation literature suggests that a prescriptive approach to scoring, which notates gestures or movement processes rather than encoding the resulting sounds, should be more amenable to working with complex timbres vis-à-vis DMIs (Kanno, 2007; Magnusson, 2019). In contrast, a descriptive approach requires that participants represent complex timbres directly, relying on the reader's interpretation to infer the desired musical outcomes.

Fourth-wave or entangled HCI scholarship complicates this binary by emphasising how users, artefacts, and cultural meanings co-constitute one another in situated practice (Frauenberger, 2019). Our findings resonate with this view: the prescriptive/descriptive manipulation did not yield isolated behaviours; instead it produced an intertwined set of abstraction/specificity and structural/metaphorical characteristics. In this sense, each notation is an artefact comprising performer gesture, sonic outcome, and visual metaphor.

In our study, we explored the process of interacting with and notating graphic scores for a novel movement-based DMI driven by a neural network. While we observed group-level differences in creative processes, conceptualisation, and direct visual-to-musical mappings, our analysis of participants' scores revealed additional, more explanatory dimensions of variation that extended beyond the prescriptive/descriptive continuum. Based on these findings, we propose a “Space of Notational Strategies,” which characterises the fundamental decisions faced when notating music. This design space spans ambiguous vs. specific and structured vs. metaphorical notation strategies, both of which apply to prescriptive and descriptive approaches alike.



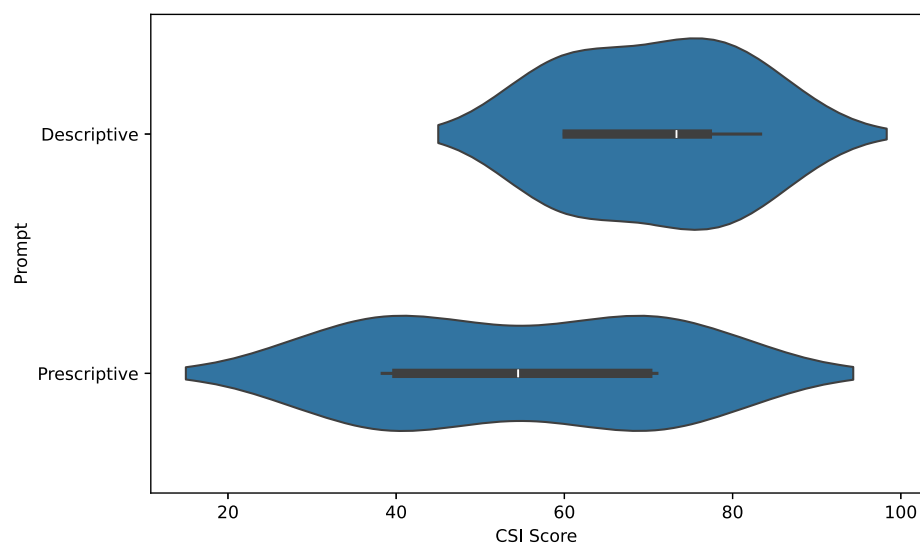


FIGURE 7
CSI score distribution by prompt.

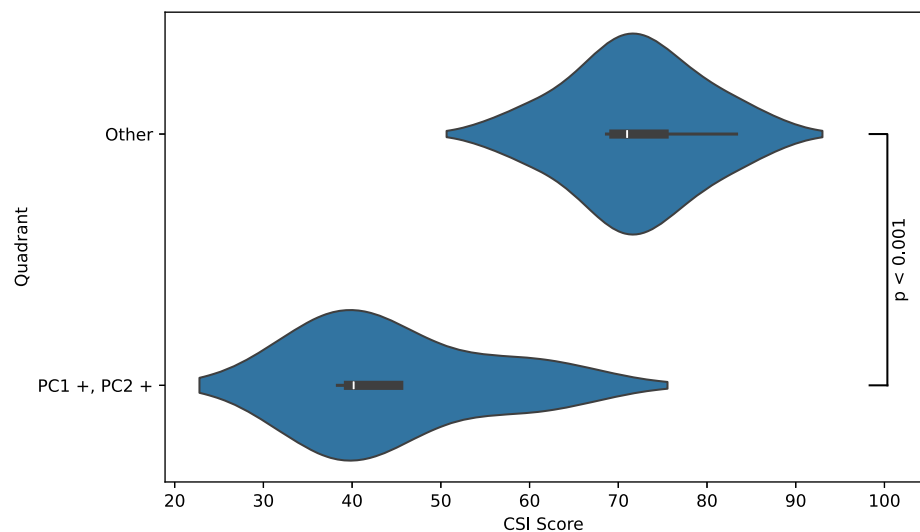


FIGURE 8
CSI score distribution by PCA quadrant. Distribution of CSI scores for participants whose graphical scores had positive PC1 and PC2 are in the lower row, those with either PC1 or PC2 negative are grouped in the upper row.

Furthermore, taken together, these dimensions appear to explain variations in participants' Creativity Support Index scores. Overall, our results underscore the diversity of approaches to graphic notation and their potential to support creative interactions with novel musical instruments.

3.3.3.1 Emergence of new dimensions

Two key dimensions of variation emerged from the PCA of the grounded visual taxonomy (see [Appendix](#) Section 1.1 for significantly loading factors):

1. A first dimension ranging from gestural and temporal specificity on one end to ambiguity, interpretation, and abstraction on the other;
2. And a second dimension ranging from musical structure and organisation (e.g., repetition, contrast, sudden changes) to metaphor and imagery.

We propose associating each dimension with a choice and a possible “strategy” for notation (see [Figure 9](#)). By considering the intersection of these binary choices we propose a four-quadrant “Space of Notational Strategies” (SNS) (see

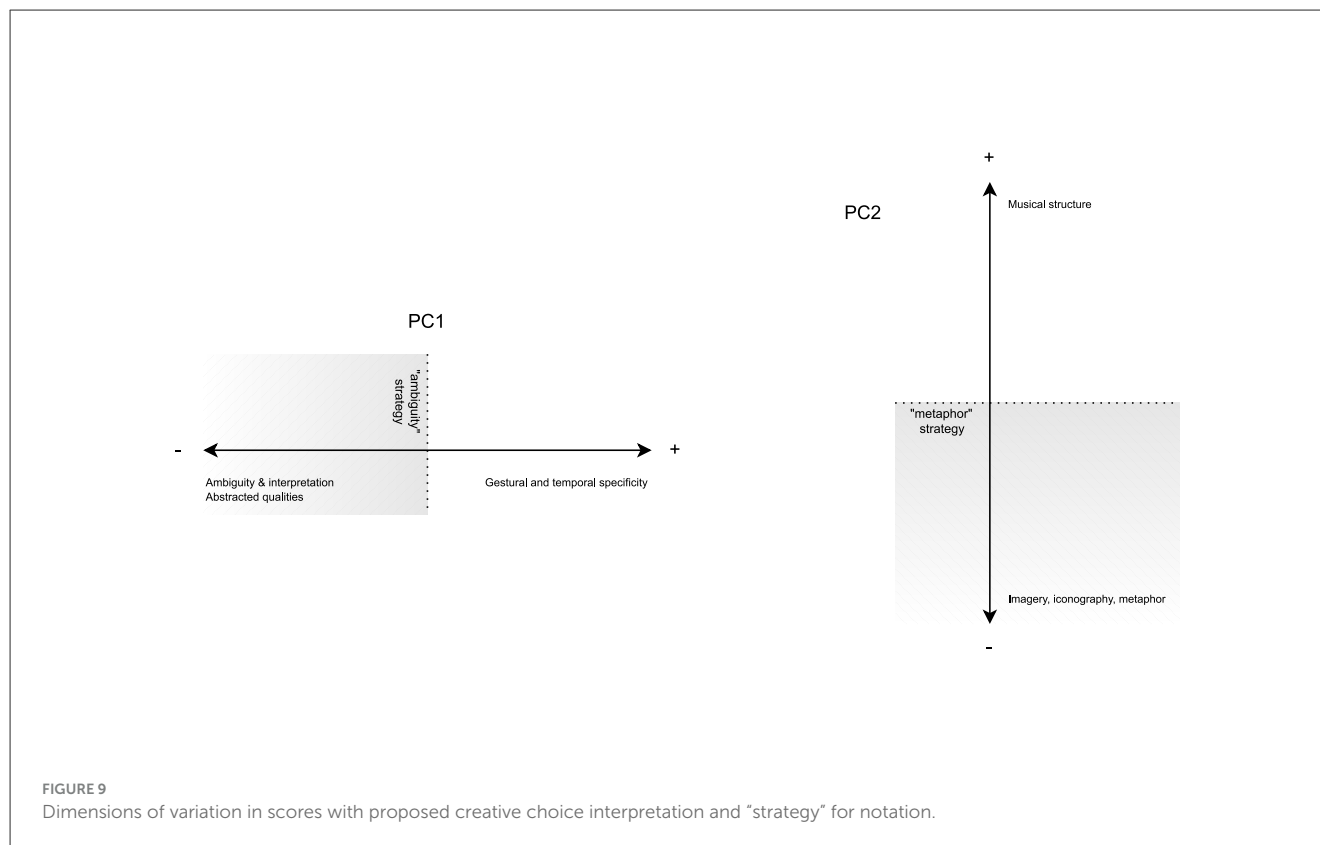


Figure 10) with implications for creative (RQ1) and notational design (RQ2) processes.

We would like now to further examine these dimensions of variation and the resulting "Space of Notational Strategies." We suggest that the dimensions emerging from our analysis reflect choices exterior to the prescriptive/descriptive divide which face a would-be composer, as well as practical "strategies" available for a notation/composition task. Our results demonstrate the use of these strategies across prescriptive and descriptive approaches, and their application is associated with the level of satisfaction reported by the composers in the creative task. Here we examine the approaches taken by clusters of participants, the visual character and graphical organisation of their scores, and the interaction between strategy, creative output, and user sentiment.

3.3.3.2 The abstraction/specificity continuum

The first principal component (PC1) of our bottom-up graphical score analysis reveals a continuum of variation ranging from gestural and temporal specificity to ambiguity and abstraction. On the positive end of this spectrum, scores exhibited a high degree of specificity, using elements such as visual alignment, staves, and structured visual guides to convey precise performance instructions. Figure 11 shows P6's score (scoring the highest on PC1) which demonstrates the specificity of gesture, temporal placement, structured visual guides, and the corresponding lack of ambiguity correlated with positive PC1.

Conversely, at the negative end of the spectrum, scores embrace ambiguity and abstraction, requiring greater interpretative input from the performer. Rather than defining every parameter

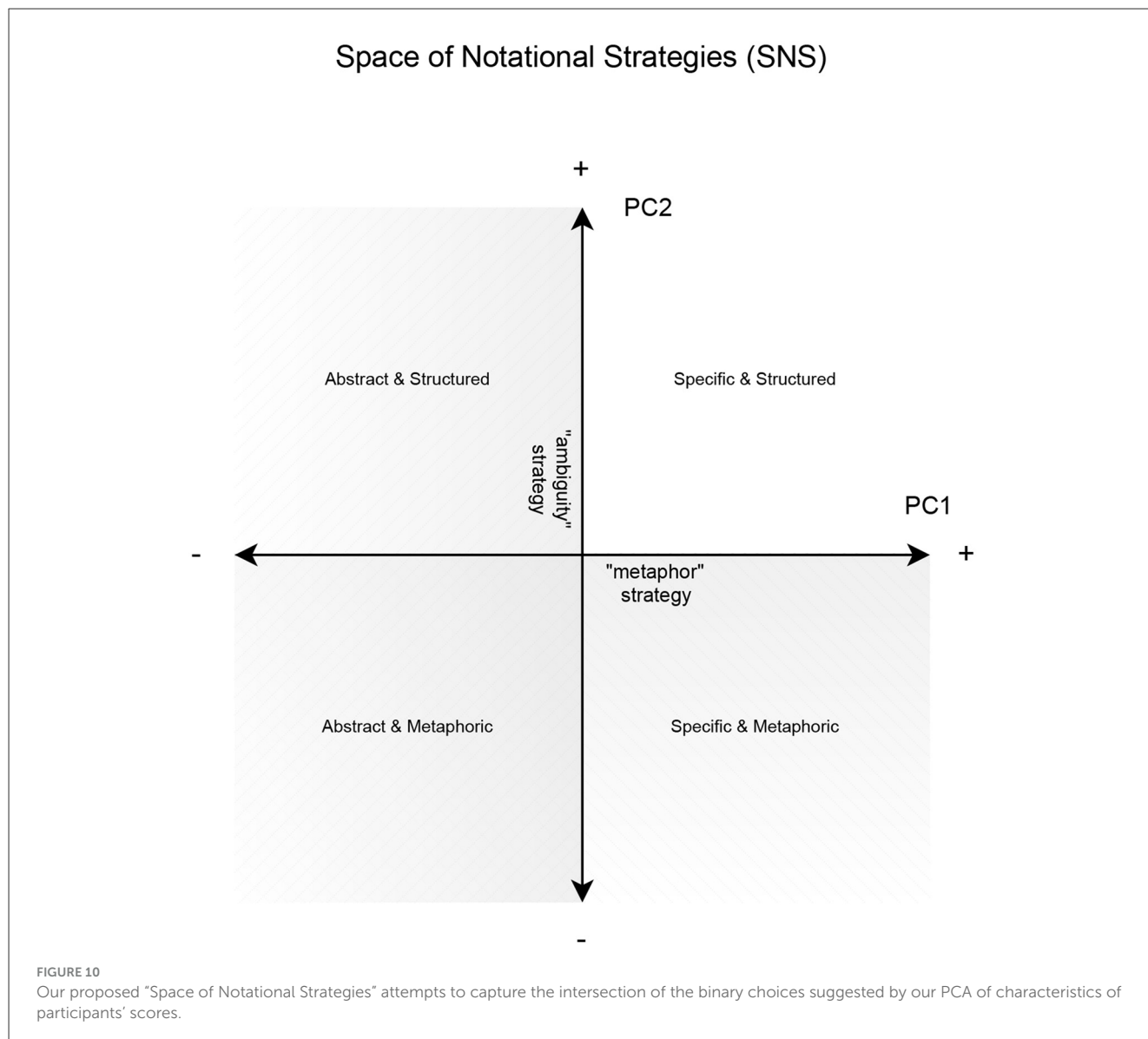
explicitly, these notations intentionally leave elements open-ended (see Figure 4 with low PC1). Musical parameters may be abstracted, so that more general qualities are described, thus requiring the performer to reconstruct details during preparation or performance. Alternatively, the composer may relinquish control over decisions such as section order or duration, shifting interpretative agency to the performer. Composers define and refer to abstracted sets of somatic qualities rather than specific movements (see for example Figure 13) (RQ1).

This strategy represents a departure from traditional, explicit notation by deliberately incorporating interpretative space. Rather than attempting to fully define the musical outcome, composers employing this approach embrace ambiguity as a notational principle. This can be seen as a distinct "ambiguity strategy," one of two primary strategies we will identify in this study. Parallels can be drawn to broader movements in contemporary music, where composers deliberately challenge the traditional expectation of fully determined scores (Thomson, 1983).

3.3.3.3 A continuum from musical structure to metaphor

The second principal component (PC2) identified in our analysis contrasts another two distinct approaches to notation: at the positive end, a focus on explicitly defining musical structure through elements such as repetition, juxtaposition, and sudden changes; at the negative end, the use of metaphor, pictography, and naturalistic references as guiding principles for performance (RQ2).

Scores that score highly on PC2 rely on structural clarity to shape musical events, avoiding metaphorical or pictorial elements. A prime example is P2's score seen in Figure 12, which



exemplifies this approach by structuring the music through precise placement, repetition, and stark juxtapositions, without invoking extra-musical imagery.

Conversely, at the low end of PC2, scores forego direct structural definition in favour of metaphor and imagery to shape interpretation. Rather than prescribing musical events in detail, these notations provide evocative references intended to inspire particular qualities of performance. Exemplifying this approach with the lowest PC2 is P4 seen in Figure 13, whose entire score functions as a single intricate metaphor, guiding the performer’s approach rather than specifying musical structures explicitly.

This strategy, which we term the “metaphor strategy,” relies on shared cultural knowledge to condense complex musical ideas into a more intuitive, immediate form. By invoking familiar images, emotions, or concepts, the composer not only expedites communication but also provides a richer contextualisation of the musical material. This strategy enables the performer to engage with the score through associative and interpretative means,

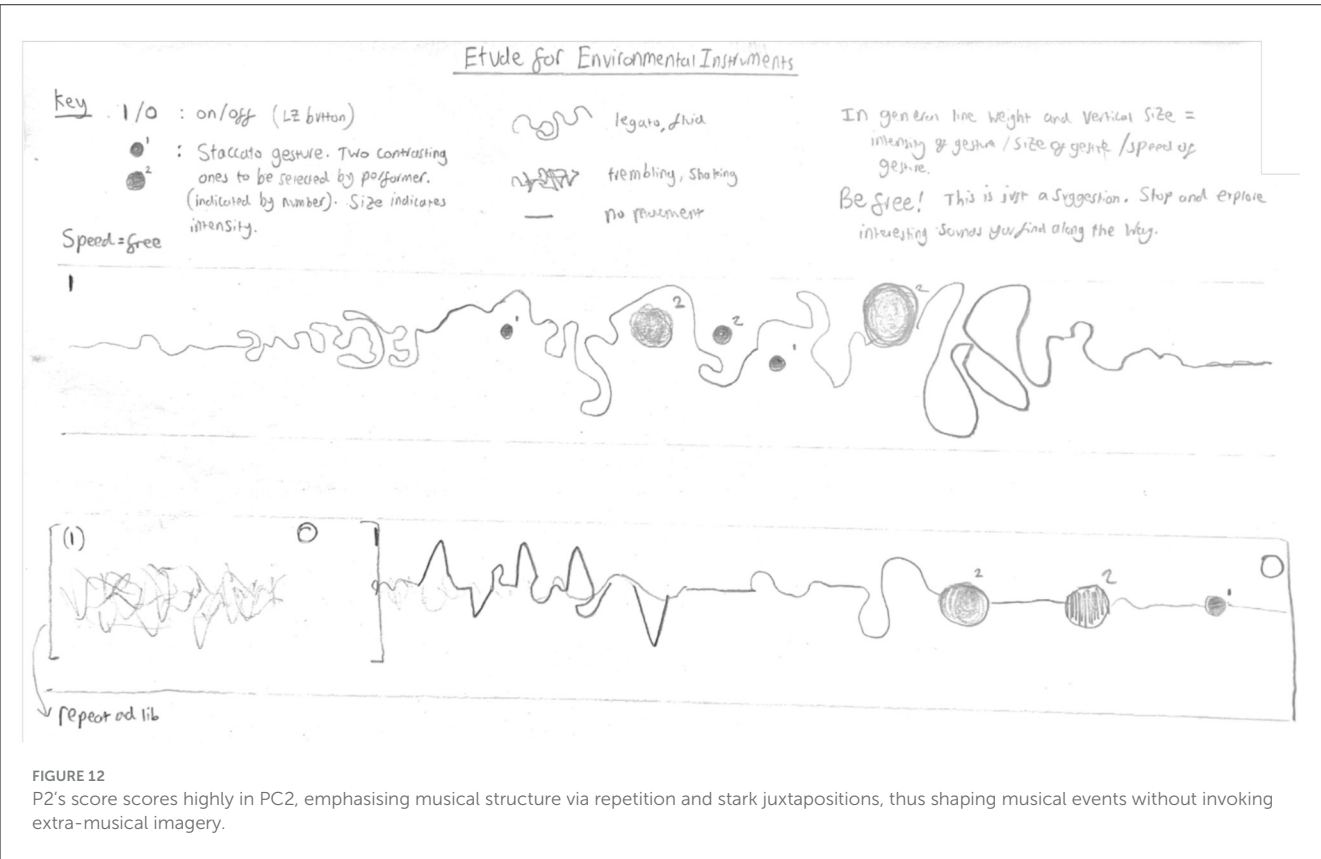
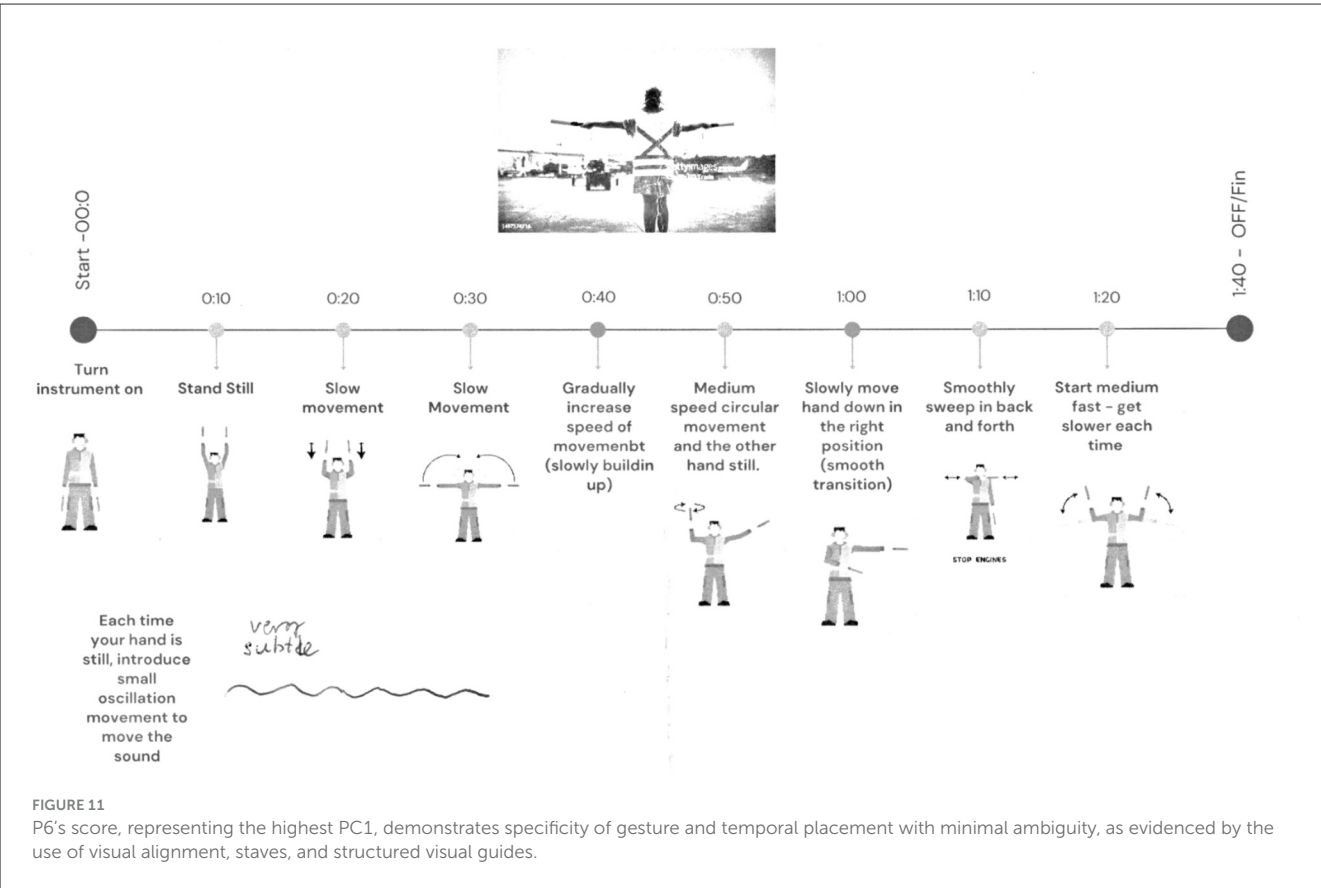
drawing on cultural and embodied understandings rather than adhering strictly to notational precision (RQ1).

3.3.3.4 Space of notational strategies

By considering the two binary choices suggested by our PCA, we propose a four-quadrant framework that we term the “Space of Notational Strategies” (SNS) (Figure 10). This model allows us to explore the relationships between different approaches to notation, analyse representative examples from our participants, and consider the broader implications for compositional practice.

3.3.3.5 Specific and structured

The SNS’s upper right quadrant (PC1 positive, PC2 positive) (see Figure 10) encompasses scores that are both highly specific in their depiction of performance actions and structured in their approach to musical organisation. Examples include the works of P1 (Figure 14), P2 (Figure 12), and P3 (Figure 14), which prioritise precise notation of events and musical structure. These scores can



be understood as the most direct or explicit response to the task of notation, emphasising clarity and unambiguous instruction.

Participants who were given the prescriptive prompt predominantly fell into this quadrant (see Figure 6), suggesting that being instructed to “notate actions” naturally led them to focus on

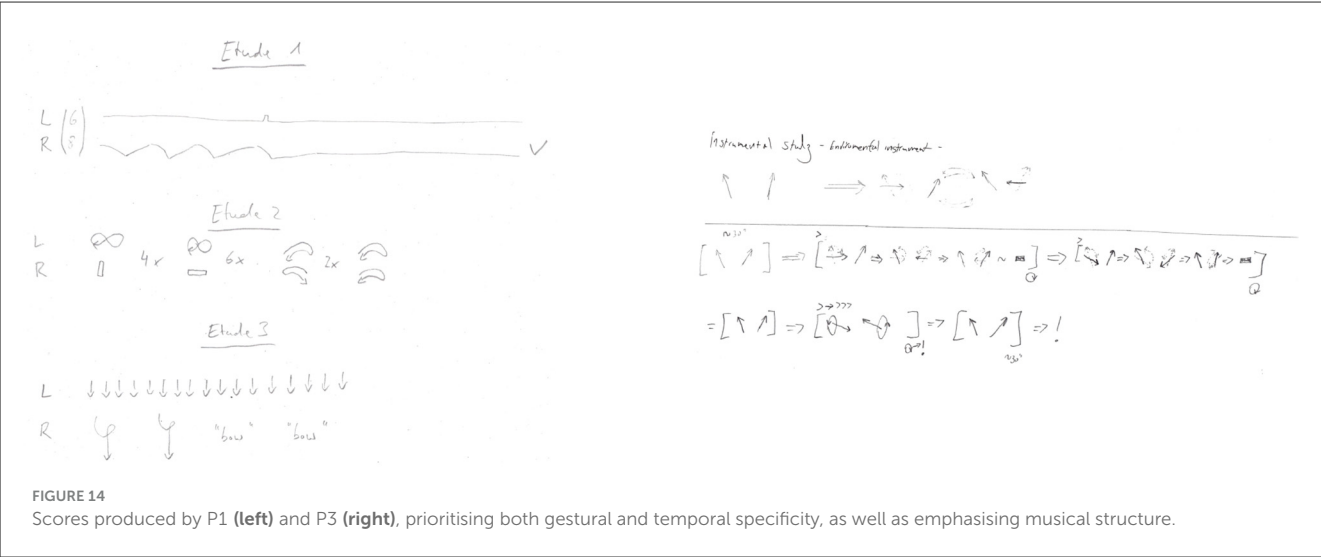
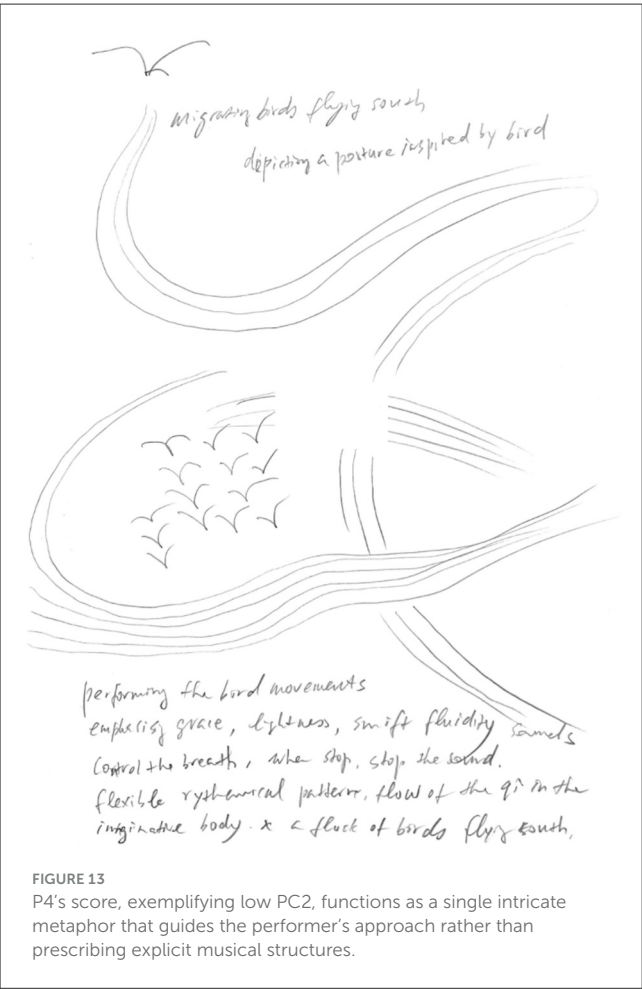
explicit and detailed descriptions of performance processes (RQ2). The prompt itself emphasised the communication of methods and actions required for performance, directing participants towards a structured and instructional approach. However, it is notable that scores in this category corresponded to the lowest Creativity Support Index (CSI) scores. This suggests that while prescriptive notation provides a rigorous means of representing complex sound and timbral structures, an insistence on specificity and structure may impose creative constraints, making the process feel restrictive or unfulfilling for composers (RQ1). Scores in this quadrant reflect the challenges participants faced in developing a precise and effective graphical language for their musical ideas.

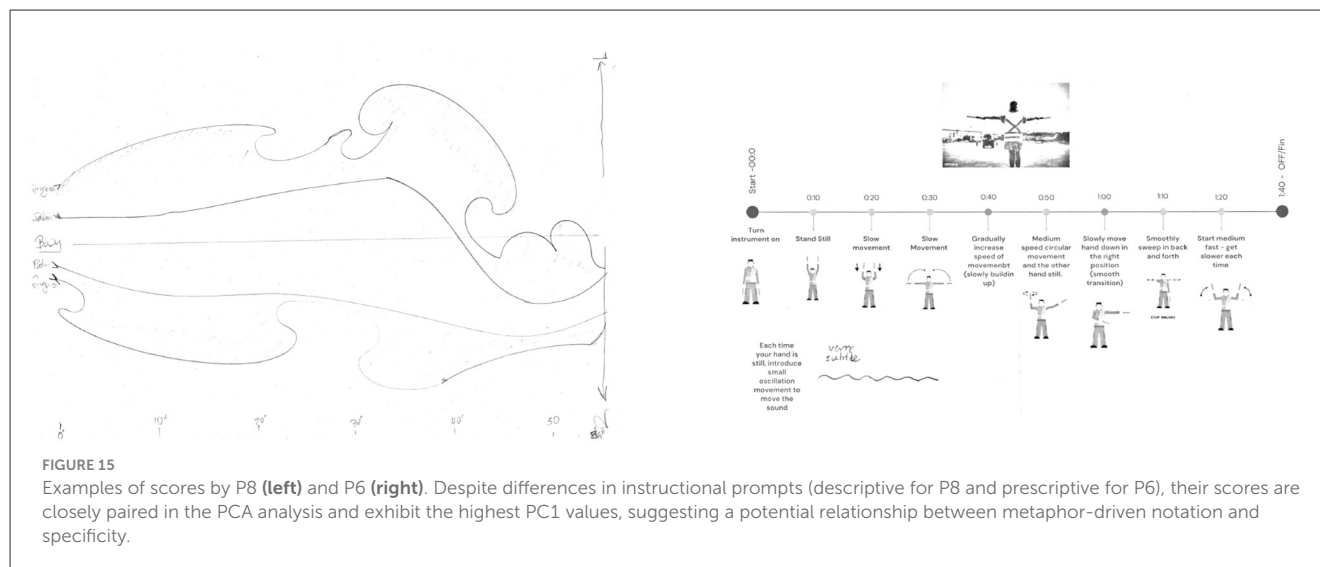
Despite the prevalence of this approach among participants in the prescriptive group, it is not the only possible response to the prompt; each quadrant of the SNS contains representatives of the prescriptive group (RQ2). While such scores may have felt restrictive to composers, as indicated by their lower CSI scores, they may offer advantages for performers by providing clear, reproducible instructions that minimise interpretative ambiguity (see Section 5.2).

3.3.3.6 Specific and metaphoric

This quadrant represents scores that apply the “metaphor” strategy while maintaining a high degree of specificity in temporal or gestural detail. Positioned at the negative end of the musical structure/metaphor continuum (negative PC2), these scores rely on evocative imagery or metaphorical frameworks to shape interpretation, yet they remain explicit in defining timing and performance actions (RQ2).

Examples include the scores of P8 and P6 (both seen in Figure 15), who were given descriptive and prescriptive prompts, respectively. Despite this difference in instruction, their scores are closely paired in the PCA analysis, suggesting a strong relationship between metaphor-driven notation and specificity. Notably, these scores exhibit the highest PC1 values, indicating that composers who adopt a metaphorical strategy may also feel empowered to be highly specific in their notation.





A striking feature of this quadrant is the prominence of explicit time organisation. Both scores in this category incorporate precise timelines, with second-by-second specifications that dictate temporal structure in detail.

3.3.3.7 Abstract and metaphoric

This quadrant represents scores that apply both the ambiguity and metaphor strategies, prioritising evocative imagery and open-ended interpretation over explicit musical structuring (RQ2). Positioned at the negative end of PC1 (indicating abstraction) and PC2 (favouring metaphor over structural organisation), these scores employ extensive graphical elements and illustrative content to guide performance (RQ1).

Two closely paired examples in the PCA analysis are P4 (prescriptive group) and P10 (descriptive group). P4's score (Figure 13) was inspired by the view from the composer's high-rise apartment, capturing the movement of birds and the wind as they migrated on an autumn evening. Rather than specifying precise gestures for the performer to execute, the score functions as a choreographic tool, using metaphor and abstraction to inspire performer movement.

In contrast, P10's score (Figure 4) employs richly illustrated iconography and symbols to represent sounds in a vivid and direct manner. The score features imagery ranging from whirling tornadoes to squealing pigs, visually conveying sonic qualities without prescribing specific techniques.

Despite differences in their structural detail and degree of prescriptive instruction, both scores demonstrate the application of metaphoric and ambiguity strategies across different notational approaches (RQ1). Instead of defining musical actions explicitly, these works rely on visual metaphor and abstract representation to shape interpretation, showing that these strategies can be effectively employed in both prescriptive and descriptive approaches.

3.3.3.8 Notational approaches and creativity support

Our findings indicate a clear relationship between notational approach and perceived creativity support. Participants who fell into the specific and structured quadrant reported significantly

lower Creativity Support Index (CSI) scores than those in all other quadrants (see Section 3.3.2.1). This suggests that a rigidly specific, structured approach to notation, while precise, may feel constrictive or creatively limiting for composers (RQ1). These observations align with existing research suggesting that divergent and convergent tasks can affect a participant's mood positively and negatively respectively (Akbari Chermahini and Hommel, 2012).

In contrast, participants who applied one of the strategies we identified (whether the metaphor strategy or the ambiguity strategy) reported a greater sense of creativity support. Engaging with these strategies may have provided participants with a sense of agency, enabling them to shape their notation in a way that resonated with their artistic intentions rather than feeling bound by conventional expectations (RQ1).

While the prescriptive prompt often led participants towards a specific and structured approach, our analysis shows that this was not an inevitability. Composers working within a prescriptive paradigm were not confined to a rigid framework unless they chose to be: alternative metaphor-driven and abstract strategies were also possible within a prescriptive context (RQ1). This reinforces the conclusion that the strategies we identified are not exclusive to either prescriptive or descriptive approaches, but can be effectively deployed across both.

3.3.3.9 Implications of a "blank slate" for embodiment and choreographic approaches

The "blank slate" character of the novel instrument we employed (its lack of established techniques, timbres, and performance traditions) appears to have encouraged embodied and choreographic thinking in participants' compositional processes, shedding light on RQ3. Several composers responded by building visual vocabularies of icons, pictograms, and illustrative scenes that linked unfamiliar sounds to familiar imagery. By grounding their ideas in recognisable symbols such as vortices, birds in flight, or gestural silhouettes, they used shared references to navigate an otherwise uncharted sound world.

The same absence of instrumental expectations also prompted a turn towards bodily movement. Unable to lean on existing playing

techniques or timbral “defaults,” some participants specified sequences of motion, posture, and spatial orientation, treating the performer’s body as primary material. In this way, choreographic thinking provided an alternative path to musical structure when no familiar instrumental behaviour was available.

These observations connect our findings on metaphor and ambiguity strategies to the broader theme of embodiment in creative work. Because such strategies can operate comfortably inside a prescriptive framework, notation is not limited to detailing discrete musical actions; it can also invite wider embodied and interpretative processes. This suggests that composers may use notation to mediate between explicit instruction and open-ended performance engagement, a perspective with important implications for embodied approaches to composition.

4 Limitations

The findings of this study must be considered within the context of several limitations. Firstly, our interpretative framework and resulting SNS are derived from a relatively small sample size of 11 participants. This limited sample reduces our ability to generalise the results. The specific structure of the SNS may thus be simplified or constrained due to this low sample size; additional dimensions or nuances might emerge in studies involving larger participant pools.

Secondly, we purposely issued each study participant with the identical NN-based instrument version. This allowed us to remove variation that would have arisen from differing versions of the instrument, but the study findings should be considered in light of having been derived in the context of a single instrument, and stand to be validated in future work with other instruments.

Thirdly, the qualitative hands-on nature of our study meant that researchers were continually aware of participants’ group assignments and individual experiences during data collection and analysis. This awareness introduces a potential source of interpretative bias, as researchers’ knowledge of group assignments could influence qualitative coding and thematic interpretation. Efforts were made to mitigate these biases through collaborative analysis and reflexive thematic methods, yet they cannot be properly ruled out.

Fourthly, although all participants were given access to the instrument at least 1 week before their focus group session, we did not prescribe or monitor the exact amount of time they spent engaging with it. This likely varied between individuals depending on availability and working style, and may have influenced the depth of their familiarity with the instrument and the kinds of creative outcomes they produced.

Fifthly, we consider the interpretative framework presented here primarily as a preliminary offering with potential to support creative approaches in music composition involving novel DMIs. The exploratory scope of this study means that its observations and theoretical proposals warrant further empirical validation. Future studies should seek to confirm whether these observations hold across broader participant groups, various DMIs, and alternative musical contexts.

Lastly, due to the study’s exploratory scope and limited dataset, further work is necessary to determine the robustness of the SNS

dimensions identified here. Future research could clarify whether the interpretative space, particularly regarding the balance between specificity and metaphor, remains consistent when applied across larger and more diverse datasets.

5 Future work

Our study provides insights into how musicians approach notation for a NN-based digital instrument, but several open questions remain. Future research should expand on our findings by refining the Space of Notational Strategies, testing notation with performers, and exploring broader theoretical implications for notation in AI-driven music.

5.1 Refining the space of notational strategies

Due to our study’s small sample size and exploratory nature, we point out that the “Space of Notational Strategies” (SNS) we propose may actually need to encompass additional dimensions that were not discernible in our analysis. We suggest that future research should reproduce our methodology with a larger participant pool and examine whether the analysis reveals further dimensions.

5.2 Evaluating interpretability with performers

Future research should investigate how performers interpret different notational strategies, particularly those generated using the SNS model. A structured experiment could task composers with creating works that exemplify each SNS quadrant. Alternatively, they could explore a broader space of notation by composing across all eight possible categories, defined by *[prescriptive vs. descriptive] × [SNS quadrants]*. This approach would help evaluate how notational constraints shape compositional behaviour and influence creativity support. Assessing the usability and effectiveness of these notations could involve:

- Creativity Support Index (CSI) ratings—measuring performers’ perceived level of creativity support,
- Thematic analysis (TA)—identifying qualitative trends in usability and interpretative strategies, and
- Convergent vs. divergent thinking analysis—examining whether different notation styles encourage structured (convergent) or exploratory (divergent) performance approaches.

5.3 Instrumental and gestural considerations

Bimanuality and bifurcation in controller-based gestural input should be further analysed to determine whether hardware

constraints influence notation paradigms. Understanding these relationships could inform the design of future gestural DMIs and their corresponding notational systems.

5.4 Perceptual graphics and structuring time

A broader question arising from our findings is whether the organisation of time is an inherent constraint within musical notation. Traditional notation systems often impose a linear, rigid temporal structure—one that may not align with the dynamic and evolving nature of AI-generated or electronic music.

As discussed in Section 3.2.2, our observations of differing visual channel usage across prompt groups point to the potential value of perceptual graphics in addressing these distinctions. The field of perceptual graphics, which investigates how visual encoding influences cognition, interpretation, and usability, offers useful perspectives for further investigating our findings. Research in this area suggests that the way information is visually represented significantly affects how it is processed and understood (Ware, 2012). Foundational studies on graphical perception and information visualisation (Cleveland and McGill, 1984) provide relevant frameworks, and work by Card et al. (1999) highlights how visual encoding strategies shape user interaction and interpretation.

Graphical notations, including animated visuals or data-driven representations, may enable more intuitive real-time engagement with musical structures. As such, perceptual graphics could serve as a complementary layer to traditional notation, or even function as a standalone system for genres that resist conventional temporal hierarchies.

These directions may ultimately broaden our understanding of notation within the context of emerging DMIs and evolving musical aesthetics.

6 Conclusion

The tripartite analyses (thematic, NDS, and grounded visual taxonomy) together clarified how prescriptive and descriptive notation shape creative interaction with a neural network-based digital musical instrument.

The results of our paper suggest that in terms of user interaction, embodiment, and creative processes (RQ1), prompt framing can lead to different notational strategies; where prescriptive prompts encouraged explicit, structured depictions of movement and action, and where descriptive prompts invited metaphor, sonic imagination, and richer semantic associations. The NDS analysis quantified these differences in mapping behaviour and feature use, making the distinction between directive specificity and more open-ended conceptual framing visible.

In terms of design and complexity of scores (RQ2), prescriptive approaches foregrounded physicality and controlled performer action, while descriptive approaches distributed interpretive agency, using abstraction and imagery to create space for embodied sense making.

In addition, our analysis of the grounded visual taxonomy revealed a latent structure underlying score variation along two core dimensions: specificity vs. abstraction and musical structure vs. metaphor. These dimensions informed our proposed Space of Notational Strategies, which captures how composers navigate trade-offs and combine strategies beyond the simple prescriptive/descriptive divide. The association between quadrant membership in that space and reported creativity support ties notational choice to user experience.

In the absence of prior instrumental expectations (RQ3), participants engaged with the novel DMI through metaphor, imagery, and embodied exploration, developing gestural strategies and metaphor-led approaches. Visual and choreographic references provided a way to connect unfamiliar sounds with familiar symbols, allowing participants to navigate the compositional task without reliance on an established sonic language.

Viewed through a fourth-wave or entangled HCI perspective, these findings indicate that DMIs, their users, and the graphic scores that mediate them form a technological and cultural mesh of mutual influence. The fact that each score can be placed independently on abstraction/specificity and structural/metaphorical continua illustrates such entanglements. The graphic score itself could perhaps then function as a point of partial “detanglement”; decisions that were previously implicit in playing are externalised in the score. By making intention and agency visible and negotiable, the score becomes a boundary object that both reveals and manages the dynamic relationship of human and technology. This observation aligns with fourth-wave HCI’s call for designs that expose and support the fluctuating relationships amongst people, creative artefacts, and cultural meanings (Frauenberger, 2019).

This work contributes a combined conceptual and empirical framework for understanding notation in novel DMIs. It provides a way to read notational artefacts, to trace how prompt design influences strategy, and to surface the latent choices available to composers. Future work should refine the Space of Notational Strategies, test its generality with more diverse instruments and users, examine performer interpretation of these notational varieties, and explore design interventions in notation systems that deliberately position users along or between the identified dimensions. Limitations include the sample size and the reliance on a specific prompt framing, both of which should be addressed in follow-up studies to strengthen the model’s robustness.

Data availability statement

The raw data supporting the conclusions of this article has been made available and is referenced as appropriate throughout the text. [Appendix Data](#) (PC loadings and participant scores) can be freely accessed here: <https://osf.io/c5fq9/>.

Ethics statement

The studies involving humans were approved by TU Berlin Kommission für Forschungsethik der Fakultät I. The studies were

conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

BB: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. LA: Conceptualization, Formal analysis, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing. HM: Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was funded in part by the Bosch Research Foundation in the Stifterverband and the Erhard Höpfner Foundation. We thank the Hasso Plattner Institute for providing compute resources to train the RAVE model used in the instrument. Open access publication costs were covered by the University of Potsdam.

Acknowledgments

The authors thank Dr. Henrik von Coler (Georgia Tech University) and Dr. Marcelo Wanderley (McGill University) for their insights and feedback throughout the course of this study.

References

- Ackva, V., and Schulz, F. (2024). “ANIRA: an architecture for neural network inference in real-time audio applications,” in *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)* (Erlangen: IEEE), 1–10. doi: 10.1109/IS262782.2024.10704099
- Akbari Chermahini, S., and Hommel, B. (2012). Creative mood swings: divergent and convergent thinking affect mood in opposite ways. *Psychol. Res.* 76, 634–640. doi: 10.1007/s00426-011-0358-z
- Auri, L., and von Coler, H. (2025). “Neural networks as catalysts for new musical cultures: a study on instrument design,” in *First International Conference on Digital Science in Culture, Potsdam, Germany* (Cham: Springer Nature).
- Bacon, B. (2022). “Rethinking the notation design space,” in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR 2022)*, eds. T. Vincent, J. Bell, and C. de Paiva Santana (Marseille, France: PRISM Laboratory), 70–77.
- Bai, S., Kolter, J. Z., and Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv [preprint]*. arXiv:1803.01271 [cs]. doi: 10.48550/arXiv.1803.01271
- Bent, I. D. (2025). *Musical Notation - Evolution of Western Staff Notation*. Chicago, IL: Encyclopedia Britannica.
- Bent, I. D., Hughes, D. W., Provine, R. C., Rastall, R., Kilmer, A., Hiley, D., et al. (2001). *Notation*. Grove Music Online. doi: 10.1093/gmo/9781561592630.article.20114
- Blackler, A., Zelenko, O., Chamorro-Koc, M., Rittenbruch, M., and Sade, G. (2018). Design “Research: methodological innovation through messiness,” in *New Directions in Third Wave Human-Computer Interaction: Volume 2 - Methodologies*, eds. M. Filimowicz, and V. Tzankova (Cham: Springer International Publishing), 131–154. Series Title: Human-Computer Interaction Series. doi: 10.1007/978-3-319-73374-6_8
- Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806
- Breen, R. L. (2006). A practical guide to focus-group research. *J. Geog. High. Educ.* 30, 463–475. doi: 10.1080/03098260600927575
- Briot, J.-P., Hadjeres, G., and Pachet, F.-D. (2020). *Deep Learning Techniques for Music Generation. Computational Synthesis and Creative Systems*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-70163-9
- Buxton, B. (2007). *Sketching User Experiences: Getting the Design Right and the Right Design*. San Francisco, CA: Morgan Kaufmann Publishers Inc.
- Cage, J., and Jandl, E. (1961). *Silence*. Middletown, CT.
- Caillon, A., and Esling, P. (2021). RAVE: a variational autoencoder for fast and high-quality neural audio synthesis. *arXiv [preprint]*. arXiv:2111.05011 [cs]. doi: 10.48550/arXiv.2111.05011
- Card, S., Mackinlay, J., and Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. Burlington, MA: Morgan Kaufmann.
- Charmaz, K. (2014). *Constructing Grounded Theory*, 2nd Edn. London: SAGE Publications
- Cherry, E., and Latulipe, C. (2014). Quantifying the creativity support of digital tools through the creativity support index. *ACM Trans. Comput. Hum. Interact.* 21, 1–25. doi: 10.1145/2617588

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. Generative AI grammar correction supplied by Overleaf (Digital Science UK Limited) was used in this paper.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomp.2025.1597806/full#supplementary-material>

- Cleveland, W. S., and McGill, R. (1984). Graphical perception: theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 79, 531–554. doi: 10.1080/01621459.1984.10478080
- Cook, N. (2013). *Beyond the Score: Music as Performance*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199357406.001.0001
- Cook, P. R. (2009). *Re-Designing Principles for Computer Music Controllers: A Case Study of SqueezeVox Maggie*. Pittsburgh, PA, 218–221.
- Corbin, J., and Strauss, A. (2015). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. London: SAGE Publications.
- Cornell Lab of Ornithology (2020). *Cornell Birdcall Identification*. Available online at: <https://www.birds.cornell.edu/home> (Accessed June 27, 2024).
- Edmonds, J. (1965). Paths, trees, and flowers. *Can. J. Math.* 17, 449–467. doi: 10.4153/CJM-1965-045-4
- Eisenhardt, K. M. (1989). Building theories from case study research. *Acad. Manag. Rev.* 14, 532–550. doi: 10.2307/258557
- Eldridge, A., and Kiefer, C. (2017). “Self-resonating feedback cello: Interfacing gestural and generative processes in improvised performance,” in *Proceeding International Conference on New Interfaces for Musical Expression (NIME 2017)* (Copenhagen), 25–29. Available online at: <https://zenodo.org/record/1176157>
- Engel, J., Resnick, C., Roberts, A., Dieleman, S., Norouzi, M., Eck, D., and Simonyan, K. (2017). “Neural audio synthesis of musical notes with WaveNet autoencoders,” in *Proceeding International Conference on Machine Learning (ICML 2017)* (Sydney, NSW: PMLR), 1068–1077. Available online at: <https://proceedings.mlr.press/v70/engel17a.html>
- Fereday, J., and Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int. J. Qual. Methods* 5, 80–92. doi: 10.1177/160940690600500107
- Frame, C. (2023). “Notating experiences: a new system for visual documentation in instrument design,” in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)* (Boston, MA).
- Frauenberger, C. (2019). Entanglement HCI the next wave? *ACM Trans. Comput.-Hum. Interact.* 27, 1–27. doi: 10.1145/3364998
- Gerhardstein, R. C. (2002). “The historical roots and development of audiation: a process for musical understanding,” in *Musical Understanding: Perspectives in Theory and Practice*, eds. B. Hanley, and T. W. Goolsby (Scarborough, ON: Canadian Music Educators’ Association).
- Godoy, R. I., and Leman, M. (2010). *Musical Gestures: Sound, Movement, and Meaning*. London: Taylor & Francis.
- Goodman, N. (1976). *Languages of Art: An Approach to a Theory of Symbols*, 2 Edn. Indianapolis, IN: Hackett. doi: 10.5040/9781350928541
- Gould, E. (2016). *Behind Bars: The Definitive Guide to Music Notation*. London: Faber Music.
- Greenbaum, T. L. (2000). *Moderating Focus Groups: A Practical Guide for Group Facilitation*. Thousand Oaks, CA: Sage Publications. doi: 10.4135/9781483328522
- Gurevich, M., and Muehlen, S. V. (2001). The accordiatron: a MIDI controller for interactive music. *arXiv [preprint]*. arXiv:2010.01574 [cs]. doi: 10.48550/arXiv.2010.01574
- Hope, C. (2017). Electronic scores for music: the possibilities of animated notation. *Comput. Music J.* 41, 21–35. doi: 10.1162/comj_a_00427
- Hunt, A., Dd, Y., Wanderley, M., and Paradis, M. (2002). The importance of parameter mapping in electronic instrument design. *J. New Music Res.* 32:18853. doi: 10.1076/jnmr.32.4.429.18853
- Hunt, A., and Kirk, R. (2000). “Mapping strategies for musical performance,” in *Trends in Gestural Control of Music*, eds. M. M. Wanderley, and M. Battier (Paris: IRCAM - Centre Pompidou), 231–258.
- Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. (2007). “The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces,” in *Proceedings of the 1st international conference on Tangible and embedded interaction, TEI '07* (New York, NY: Association for Computing Machinery), 139–146. doi: 10.1145/1226969.1226998
- Kanno, M. (2007). Prescriptive notation: limits and challenges. *Contemp. Music Rev.* 26, 231–254. doi: 10.1080/07494460701250890
- Krueger, R. A., and Casey, M. A. (2015). “Focus group interviewing,” in *Handbook of Practical Program Evaluation*, eds. K. E. Newcomer, H. P. Hatry, and J. S. Wholey (Hoboken, NJ: WILEY), 506–534. doi: 10.1002/9781119171386.ch20
- Lavastre, B., and Wanderley, M. M. (2021). “Interaction strategies in composition for karlax and acoustic instruments,” *International Symposium on Computer Music Multidisciplinary Research* (Cham: Springer), 163–179. doi: 10.1007/978-3-031-35382-6_13
- Leman, M., Maes, P.-J., Nijls, L., and Van Dyck, E. (2018). What is embodied music cognition? *Springer Handbook of Systematic Musicology*, ed. R. Bader (Cham: Springer), 747–760. doi: 10.1007/978-3-662-55004-5_34
- Maes, P.-J., Leman, M., Palmer, C., and Wanderley, M. (2014). Action-based effects on music perception. *Front. Psychol.* 4:1008. doi: 10.3389/fpsyg.2013.01008
- Magnusson, T. (2011). Algorithms as scores: coding live music. *Leonardo Music J.* 21, 19–23. doi: 10.1162/LMJ_a_00056
- Magnusson, T. (2019). *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions*, 1 Edn. London: Bloomsbury Academic. doi: 10.5040/9781501313899
- Mathews, M. V. (1991). The radio baton and conductor program, or: pitch, the most important and least expressive part of music. *Comput. Music J.* 15, 37–46. doi: 10.2307/3681070
- McKay, T. (2021). *A Semiotic Approach to Open Notations: Ambiguity as Opportunity*. Cambridge, MA: Cambridge University Press. doi: 10.1017/9781108884389
- McPherson, A., Morrison, L., Davison, M., and Wanderley, M. M. (2024). On mapping as a technoscientific practice in digital musical instruments. *J. New Music Res.* 53, 110–125. doi: 10.1080/09298215.2024.2442356
- McPherson, A. P., and Kim, Y. E. (2012). The problem of the second performer: building a community around an augmented piano. *Comput. Music J.* 36, 10–27. doi: 10.1162/COMJ_a_00149
- Migiro, S. O., and Magangi, B. A. (2011). Mixed methods: a review of literature and the future of the new research paradigm. *Afr. J. Bus. Manag.* 5, 3757–3764.
- Miller, M., Häußler, J., Kraus, M., Keim, D., and El-Assady, M. (2018). Analyzing visual mappings of traditional and alternative music notation. *arXiv [Preprint]*. arXiv:1810.10814. doi: 10.48550/arXiv.1810.10814
- Morreale, F., and McPherson, A. P. (2017). “Design for longevity: ongoing use of instruments from NIME 2010–14,” in *New Interfaces for Musical Expression* (Copenhagen: Aalborg University Copenhagen).
- Müllensiefen, D., Gingras, B., Musil, J., and Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS ONE* 9:e89642. doi: 10.1371/journal.pone.0089642
- Nattiez, J. (1990). *Music and Discourse: Toward a Semiology of Music*. Princeton, NJ: Princeton University Press.
- Palacio-Quintin, C. (2008). “Eight years of practice on the hyper-flute: technological and musical perspectives,” in *Proceeding International Conference on New Interfaces for Musical Expression (NIME 2008)* (Genova), 293–298. Available online at: <https://zenodo.org/record/1179609>
- Ramachandran, V. S., Marcus, Z., and Chunharas, C. (2020). “Bouba-kiki: cross-domain resonance and the origins of synesthesia, metaphor, and words in the human mind,” in *Multisensory Perception* (Amsterdam: Elsevier), 3–40. doi: 10.1016/B978-0-12-812492-5.00001-2
- Seeger, C. (1958). Prescriptive and descriptive music-writing. *Music. Q.* 44, 184–195. doi: 10.1093/mq/XLIV.2.184
- Siedenburg, K., Saitis, C., McAdams, S., Popper, A. N., and Fay, R. R. editors (2019). “Timbre: acoustics, perception, and cognition,” in *Springer Handbook of Auditory Research, Vol. 69* (Cham: Springer International Publishing). doi: 10.1007/978-3-030-14832-4
- Tanaka, A. (2019). “Embodied musical interaction: body physiology, cross modality, and sonic experience,” in *New Directions in Music and Human-Computer Interaction*, eds. S. Holland, T. Mudd, K. Wilkie-McKenna, A. McPherson, and M. M. Wanderley (Cham: Springer International Publishing), 135–154. Series Title: Springer Series on Cultural Computing. doi: 10.1007/978-3-319-92069-6_9
- Thomson, W. (1983). Functional ambiguity in musical structures. *Music Percept.* 1, 3–27. doi: 10.2307/40285248
- Waisvisz, M. (1985). “The hands: a set of remote MIDI controllers,” in *Proceeding International Computer Music Conference (ICMC 1985)* (Vancouver, BC: International Computer Music Association).
- Ware, C. (2012). *Information Visualization: Perception for Design*, 2nd Edn. Amsterdam: Elsevier.
- Wessel, D., and Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Comput. Music J.* 26, 11–22. doi: 10.1162/014892602320582945
- Wishart, T. (1996). *On Sonic Art*. London: Routledge. doi: 10.4324/9781315077895
- Wu, Y., Bryan-Kinns, N., and Zhi, J. (2022). Exploring visual stimuli as a support for novices creative engagement with digital musical interfaces. *J. Multimodal User Interfaces* 16, 343–356. doi: 10.1007/s12193-022-00393-3
- Yamaguchi, Y., Kawada, T., Nagahama, T., and Horita, T. (2022). “A pilot study of the MVP support system using Google Glass,” in *EdMedia+ Innovate Learning* [Association for the Advancement of Computing in Education (AACE)], 17–28.
- Zheng, S., Xambó Sedó, A., and Bryan-Kinns, N. (2024). A mapping strategy for interacting with latent audio synthesis using artistic materials. *arXiv [preprint]*. arXiv:2407.04379. doi: 10.48550/arXiv:2407.04379