Evaluation of musical instruments from the musician’s perspective: A questionnaire for assessing the musician’s perception of the experiential qualities of musical instruments

Master Thesis

Gian-Marco Schmid, B.Sc.
Institute of Psychology
Department for Cognitive Psychology and Methodology
University of Basel

November 2015

Thesis Supervisors:

Dr. Alexandre N. Tuch
Center for Cognitive Psychology and Methodology, University of Basel

Prof. Dr. Klaus Opwis
Center for Cognitive Psychology and Methodology, University of Basel

Abstract

Within the research field of New Interfaces for Musical Expression there is a lack of a common ground and accepted methods for musical instrument evaluation. This master thesis aims to assess this issue by developing a psychometric tool for evaluating the experiential qualities of musical instruments from the perspective of the musician. By embracing an exploratory approach, eleven interviews with musicians from various backgrounds have been conducted. Based on the interviews, an initial item pool was created and subsequently reduced in an online study with N = 47 experts in the field of music research. In a second online study the underlining factor structure of N = 75 items was investigated using exploratory factor analysis with N = 300 participants using different kinds of musical instruments. The results suggested a three factor solution. The interrelated facets of the experiential quality criteria of musical instruments are (1) experienced freedom and explorability, (2) perceived control and comfort and (3) perceived stability, sound quality and aesthetics. The emerged structure is similar to previously conducted research on the violin. This suggests that musicians using different musical instruments tend to have a similar concept of instrument quality. This is a first indication of a possible application of the developed questionnaire for different musical instruments in various use cases.

Introduction

Evaluation in the research field of New Interfaces for Musical Expression (NIME) has recently reached high attention (Barbosa, Malloch, Huot, & Wanderley, 2015). A tendency towards a musician-centered perspective can be observed by a majority of evaluation studies that investigated the performer's perspective (Barbosa et al., 2015; Morreale, Angeli, & O'Modhrain, 2014). This emphasis has apparently been present since the beginning of the NIME-conference in 2001 and may be a result of the conference's origin from the field of Human-Computer Interaction (HCI) (e.g. Wanderley & Orio, 2002). Currently, musical instrument designers and researchers in the NIME-community more and more frequently assess the user experience (UX) of newly developed instruments to acquire a holistic view of how musicians experience them (Stowell, Robertson, Bryan-Kinns, & Plumbley, 2009; Young & Murphy, 2015). This allows for improvement and guidance in the design of new instruments by detecting possible design flaws and the conceptualization of desired experiential qualities such as expressiveness, control, learnability, playability and enjoyment. This master thesis strives to find the key experiential quality criteria of musical instruments and aims to develop a questionnaire for their assessment - suitable for a broad range of musical instruments.

A pivotal element for assessing NIMEs is their distinction to other instruments. NIMEs often fall into the category of electronic and digital musical instruments (DMI). This does not exclude non-digital instruments from the research scope of the NIME-community, but with the rise of available low-cost sensors and open source audio solutions (Franinović & Serafin, 2013), DMIs have shown to be a major group of investigated instruments in NIME-research (Barbosa et al., 2015). In contradistinction to traditional musical instruments, DMIs are musical devices where the gestural interface is separable from the sound producing unit (Malloch, Birnbaum, Sinyor, & Wanderley, 2006). Since the interface and sound unit are separable, evaluation can be done on different levels of granularity by evaluating distinct elements of a whole system with different goals (Barbosa et al., 2015). For example comparing different instruments on only a target design element such as the haptic modality. Furthermore, various perspectives ranging from a designer's to an audience point of view can be considered (O'Modhrain, 2011). Altogether this led to novel challenges in NIME-evaluation. In order to account for those demands, a handful of evaluation frameworks have been introduced (Jordà & Mealla, 2014; Morreale et al., 2014; O'Modhrain, 2011) that guide designers of NIMEs by helping them in deciding which evaluation criteria could be important with reference to the possible perspectives of interest. Because NIME arose from the field of HCI (Poupyrev, Lyons, Fels, & Blaine, 2001) previous frameworks suggested to apply and adapt established HCI-methods, such as usability-testing in a musical context (Orio & Wanderley, 2002). Kiefer, Collins and Fitzpatrick (2008) reflected this perspective and criticized that usability testing of DMIs would lack to provide any measure of the musician's experience in musical interaction. Thus, rather recently, the realignment to a holistic user/musician experience perspective that includes constructs such as flow, enjoyment, engagement and emotion has been proposed (Swift, 2013). By looking at jamming, Swift argues that a musician-centered view would serve as a more suitable approach for creative domains because the experience of the action itself is in the spotlight. In contrast, the experience of accomplishing a goal aims at a specific outcome (e.g. the experience of formatting a text). This corresponds to Hassenzahl, Kekez and Burmester's (2002) notion of goal and action mode describing different mental user states.

The current situation of evaluation in NIME research was recently reviewed by Barbosa et al. (2015). By looking at the NIME-conference proceedings from 2012 to 2014 the authors observed a clear increase in interest in evaluation as such. They also found that a common definition of the term evaluation itself and a general consensus about relevant evaluation criteria has not have been reached yet, as it has been done in related fields. By looking at the essential stakeholders (performer, designer and the audience) Barbosa et al. detected that the performer's perspective is the most investigated role in NIME-evaluation. With regard to the evaluated objects, DMIs as whole instruments are the most common target of evaluative studies. Despite emerging trends in using specific quality criteria, there appears to be no consensus about the importance of the used criteria and on how these criteria should be assessed. Remarkably, even the reasoning behind the choice of specific criteria sometimes remains unclear (e.g. Marshall & Wanderley, 2011). More precisely, there is currently no standardized psychometric measurement tool available that allows a valid and reliable assessment of the subjectively experienced qualities of musical instruments from a performer's perspective. As a consequence, user studies within NIME often use self-developed ad-hoc questionnaires without indication of the psychometric properties of the used scales (Schmid, 2014). These measurement tools thus fall short in providing a valid and reliable assessment of the investigated criteria. As a result, interpretations and subsequent design decisions drawn from those measurements may be misleading. Hence, replication remains challenging and studies are difficult to compare.

Therefore, Stowell, Plumbley and Bryan-Kinns (2008) demanded for further development of evaluative methods. Although there are psychometric measurement tools available from HCI (e.g., for the evaluation of software, websites or smartphones) that could potentially be used with NIMEs (Young & Murphy, 2015), the application of those questionnaires to a musical context however would neglect the important step of initially examining experience dimensions that are specifically important for musicians when evaluating musical instruments. Furthermore, results may be difficult to interpret from psychometric tools directly overtaken from HCI because they were developed and standardized for other domains of interest.

We expect that the resulting questionnaire of this work allows instrument designers to illustrate how musicians perceive their instruments on the different quality dimension and whether this corresponds to the intention of the designer. These quality dimensions are thought to be universal to some extent, similar to psychometric tools in the field of HCI for the assessment of a broad range of interactive products (e.g. Hassenzahl, Burmester, & Koller, 2003). Hence, the questionnaire could be applied to evaluate different instruments in various contexts. For example, the tool could provide insights into benefits from haptic augmentation of DMIs; e.g. which quality dimensions are relevant when designing for the haptic modality (Papetti, Schiesser, & Fröhlich, 2015). The evolved psychometric tool will provide a better understanding of how musicians build an overall quality impression of musical instruments and thus support both designers and researchers in the development and evaluation of musical interfaces and instruments.

This master thesis aims at (1) finding the relevant experiential quality criteria from the musician's perspective (2) develop an initial psychometric measurement tool for their assessment (3) and reveal the underlying structure of the musician's perceived experiential quality criteria of musical instruments. The usage of the term 'musical instrument' in our work refers to Bongers' (2000) general notion of an interactive device where the musician can control sound-production by various physical gestures (Malloch et al., 2006). Before the detailed description of the undertaken research and its respective analysis is presented, a brief overview of the relevant background for this work follows.

Related work

This section gives a brief overview of the topic of evaluation in NIME-research by looking at two relevant evaluation frameworks in the field and often assessed evaluation criteria. After that we take a glimpse at the related field of general musical instrument evaluation. The section concludes with a synthesis of this outlook and its implications for our work.

Evaluation in the domain of New Interfaces for Musical Expression

Only a few years ago Stowell, Plumbley and Bryan-Kinns (2008) noted that there is very little research about evaluation of NIME; but lately, the evaluation of NIME became a topic of major interest (Barbosa et al., 2015). In their review of the NIME-conference proceedings between 2012 and 2014 that applied the term evaluation, Barbosa et al. looked at the considered targets that are often evaluated (DMI, Input, Mapping, Output, Feedback or Performance), stakeholders (performer, audience and designer), goals (e.g. comparison of instruments), evaluation criteria (e.g. controllability), the used research methods (e.g. interviews) and the duration of the evaluative assessments. The authors conclude that there appears to be a lack of general consensus about evaluation in NIME. This matter may be reflected by the notable amount of studies that label themselves with "evaluation", but omitted to report used criteria, methods or sometimes even goals of the conducted evaluations (Barbosa et al., 2015). This is mentioned to prevent the assessment of validity and replicability in research (Greenberg & Buxton, 2008).

Regarding the term evaluation, O'Modhrain (2011) claimed a broadening of the scope of evaluation. Since the four proposed stakeholder groups in her framework - audience, performer, designer and manufacturer - apply a different understanding of evaluation, each of them should be accounted for regarding their special interests. For example, while a manufacturer may be more interested in market surveys and sales, the performer demands an in-depth evaluation of the instrument itself. Based on this and other works, Morreale et al. (2014) presented a user-centered experience framework, which distinguishes between goals (purposes of the interface; e.g. user story) and specifications (interaction requirements derived from the goals).

Evaluation criteria in NIME

From the few studies that indicated evaluation criteria, the observation of the word cloud analysis (larger words represent more often assessed criteria) by Barbosas et al. (2015) suggests that expressiveness, control, learnability, playability, intuitiveness and fun are the core constructs of interest in NIME-evaluation of the performer's perspective. Subsequently, we briefly and not conclusively discuss the central concepts on which NIMEs were previously evaluated by looking mostly at the performer's perspective.

Juslin (2003) conceptualized the construct of performer expression in music as a multi-dimensional construct with five facets (GERMS Model). (1) Generative rules apply to the rule-based transformation of scores into music. (2) Emotional expression refers to the ability of communicating emotions through a corresponding musical play. (3) Random variability describes the aspect of performing which is not completely controlled by the musician and makes every performance unique. (4) Motion principles can be intentional for creating specific patterns (e.g. ritardando) or non-intentional by referring to physiological limitations of the body and (5) stylistic unexpectedness holds for the created tension by violating and resolving musical expectations. Poepel (2005) operationalized chosen aspects of the GERMS Model for comparing the expressivity of three string instruments (traditional and electric violas with different interface-mapping-syntheses). Jordà and Mealla (2014) assessed expressiveness for the evaluation of performances with different DMIs from the audience's perspective. Kontogeorgakopoulos and Kouroupetroglou (2011) compared the performer's expressivity in two musical tasks with and without haptic feedback using the Falcon haptic device (a haptic control interface). Erkut, Jylhä and Discioglu (2011) presented a model for the design and evaluation of musical interfaces including expression as one dimension of interest among others.

Orio and Wanderley (2002) were specifically interested in adapting usability testing to a musical context and highlighted several concepts suitable for musical tasks (e.g. pitch control or rhythm control). They suggested dividing control into feature controllability (manipulation of sound parameters) and timing controllability (how precisely a performer can play along a given tempo) for the assessment of musical tasks. Johnston, Candy and Edmonds (2008) reflected control from a musician-centered perspective by introducing three different modes of interaction that musicians can occupy. In instrumental mode, musicians aim to control the behavior of the instrument. An ornamental mode emerges when the musician is not intentionally aware of the exact outcome, but rather is positively surprised by the instrument itself. The conversational mode refers to a balance between the instrumental and the ornamental mode. In an evaluation study of the 'Viblotar' (a monochord-type DMI designed for experimental investigations), Marshall and Wanderley (2011) measured the subjectively perceived controllability and ease of use with and without haptic feedback. The subjectively perceived accuracy in sound modification was also investigated for the Falcon haptic device (Kontogeorgakopoulos & Kouroupetroglou, 2011). Birnbaum, Fiebrink, Malloch and Wanderley (2005) looked at musical control divided in three levels (timbral level, note level, and control over a musical process) in a proposed dimension space for DMIs which allows to categorize DMIs.

With relation to control the concept of learnability is extensively discussed: According to Jordà (2004) the ultimate goal for a designer should be to design instruments that are both appealing to the expert and the beginner. Jordà exemplifies this issue by outlining the learning curves of different tradition musical instruments. Whereas the kazoo (a membranophon which modifies the singing voice) can be learned fast and easily, it does not offer lifelong mastery. A violin on the other hand does, but is very hard to learn during the early years. An optimal learning curve is provided by the piano. Even a beginner could soon play beautiful melodies and still develop his musicianship over a lifetime. This notion is supported by O'Modhrain's (2011) remark that a challenging DMI would help to develop virtuosity. Also Wallis, Ingalls, Campana, and Vuong (2013) theoretically discuss which qualities of musical instruments contribute to long-term engagement by looking at self-determination theory of motivation (Ryan & Deci, 2000) and discussing the motives of mastery, autonomy and purpose. Also in that context, an optimal learning curve (Marshall and Wanderley, 2011) or adequate intuitiveness (Overholt, 2009) are expected to be related to engagement. By looking at the DMI categorization of Miranda and Wanderley (2006) where DMIs are divided into instrument-like controllers, extended instruments, instrument-inspired controllers and alternative controllers, Young and Murphy (2015) explained that in alternative controllers familiarity is actively avoided. Thus, intuitiveness may not be a generally applicable facet of interest. Furthermore, enjoyment, entertainment, fun or pleasantness are often assessed concepts with regard to assessing UX in a musical context (Kontogeorgakopoulos & Kouroupetroglou, 2011; Marshall & Wanderley, 2011). Finally, DMIs were also assessed for their explorability, degrees of freedom or their operational freedom (Birnbaum, et al., 2005), range of expression (Overholt, 2009) or expression per se (Erkut, Jylhä, & Discioglu, 2011).

This short glimpse only reflects the major concepts of interest that were previously investigated in NIME research. Because our approach is based on various instruments we likewise wanted to consider possible learnings from evaluation research on traditional musical instruments. Because of the predominant work on the violin in this area, a brief overview over the related work in that area is presented.

Evaluation of traditional musical instruments - The example of the violin

Within the general field of musical instrument evaluation investigations on traditional musical instruments, sensory perception, preference judgments, and their correlation to objective, physical measures form the dominant research paradigm (Fritz & Dubois, 2015). Fritz and Dubois (2015) reviewed the growing research body in this domain by discussing studies in the field of musical acoustics, which investigates the musical quality of instruments. Methodically, these studies focus on the experimental method of listening and playing tests where the double-blind format has emerged as the gold standard. Moreover, the general aim of these works is to scientifically define a ‘good’ or a ‘bad’ instrument. For example, a controversially discussed series of studies (Claudia Fritz et al., 2014; Claudia Fritz, Curtin, Poitevineau, Morrel-Samuels, & Tao, 2012; Wollman, Fritz, & Poitevineau, 2014) found that new violins were preferred over old Italians. Fritz et al., (2014) concluded that monetary value and historical importance might be the cause for a biased impression of violins. According to Fritz et al., (2014), subjective taste and playing qualities appear to be more important aspects that account for violin preference. The authors therefore suggested that future research should investigate how musicians subjectively evaluate musical instruments and which playing qualities of the instrument influence their evaluation. It is also mentioned that higher-level perceptual processes are involved when musicians choose instruments (e.g. in a music store).

In conclusion, in NIME-research a number of potentially important quality criteria have been discussed for their assessment of different stakeholder perspectives. However, it remains unclear which of the quality criteria in NIME-research truly contribute to or are more important than others for the subjectively perceived quality of a musical instrument from the perspective of the musician. Furthermore, the relations among the criteria and their individual relation to the higher-level construct of the perceived musical instrument quality remain unclear in both NIME-research and traditional musical instrument evaluation.

Research goals and motivation

To our knowledge there is currently no psychometric evaluation tool available for the NIME-community that measures the perceived instrument quality. This work is a first attempt towards the development of a UX-inspired psychometric tool that allows to assess experiential qualities of a musical instrument. These experiential qualities are expected to represent relevant attributes of a musical instrument, and as being perceived as relevant for the musical experience by the musician. The developed questionnaire is expected to allow musicians, designers and researchers to evaluate different facets of the subjectively perceived quality of musical instruments. Because of the intention of exploring its core dimensionality, this work follows an iterative bottom-up approach and is guided by the following goals:

1) Uncover the key experiential qualities of musical instruments from the musician's perspective.
2) Analysis of the importance of those experiential qualities from the perspective of music research experts and the agreement among them.
3) Development of a psychometric measurement tool for the assessment of the musician's perception of experiential qualities of musical instruments.

Methodological approach

The research project was divided into three consecutive studies (inspired by the development process of questionnaires in Psychology (Allen & Yen, 2001) and HCI (Moshagen & Thielsch, 2010): (1) To identify a broad spectrum of possibly relevant quality criteria, in a first stage, eleven semi-structured interviews with various musicians from different backgrounds have been conducted in order to identify potentially important criteria. The thereby collected criteria were sorted by means of affinity diagramming. Semantically redundant criteria were dropped, resulting in a first list of potential questionnaire items. (2) To further reduce and refine the item pool, all items were evaluated by experts in the field of music research (N = 47) in regard to their importance for the subjective evaluation of musical instruments. (3) In a third step, an online study where musicians (N = 300) had to use the initial version of the questionnaire was conducted to uncover the underlying factor structure.

Study 1: Interviews with musicians and item generation

Study 1 aimed at laying the basis for the subsequent studies by gathering as much as possible relevant quality criteria by interviewing musicians with various backgrounds. The interviews served as input for developing an initial item pool. Additionally, an affinity diagramming workshop was conducted in order to get a first impression of the possible core themes and for guidance in the item generation phase.

Method

Participants. Eleven musicians from various backgrounds were interviewed. All were male and on average 28.9 years (*SD* = 3.48) old. In order to gain a broad amount of information about important aspects of musical instruments, musicians playing different musical instruments and having diverse musical backgrounds were interviewed. Two of them are professional musicians with an academic background in music (a pianist and a percussionist). Four can be best described as semi-professional musicians who had an extensive musical education and are currently very active musicians in various projects (a pianist, an accordionist, an individual artist using different electrophonic instruments, and a saxophonist who also plays keyboard instruments). The remaining five musicians play a musical instrument as a recreational activity, however, still investing a reasonable amount of their time in musical practice (two guitarists, one electronic bass player, and two musicians who preferably interact with electronic musical instruments such as MPCs, sampler and application-based computer workstations mainly for the purpose of working as a DJ). An overview with detailed information about the interview participants is given in Table 1.

Procedure. All interviews followed the same guiding question: *What are the central dimensions that are relevant for the subjective quality of musical instruments?* This question was operationalized by the following procedure: First, the participants were asked to describe a situation where they tried a new musical instrument that they very much liked. By using an adapted form of interview-laddering (Reynolds & Gutman, 1988) the interviewees were, after they described the situation, asked in detail why the specific criteria they mentioned were important for them being delivered by a ‘good’ musical instrument. After that, participants were asked to report a situation where they disliked an instrument they tested. This time the interviewees were asked which criteria made a musical instrument ‘bad’ and why this was the case. In a third step, the interview shifted towards a more abstract level, where we asked the musicians to report what features belong to either a ‘good’ or a ‘bad’ musical instruments. In the final part of the interview the musicians were asked to describe a positive and negative experience in any musical context. This part also included questions about what they thought would be important for them in music-making in general, how they would describe the experience of playing a musical instrument, and similar questions related to everyday musical practice. The interviews were audio-recorded and lasted between 40 to 60 minutes. Based on the interviewer’s notes and the audio-recordings a summarized transcript of each interview was prepared for further analysis.

Results Study 1

From the interview-transcripts coherent thought-units were extracted (from interview step one to three). A thought-unit represents a contiguous statement from an interviewee. Over all interviews a total of 270 thought-units were identified (mean number of thought-units per interview = 24.55, *SD* = 7.55). Next, each thought-unit was reviewed for a further division in coherent subunits because observations showed that different quality criteria were mentioned in the same sentence. In total 330 single statements (mean = 30.00, *SD* = 9.58) resulted from that step. In order to reduce this set of subunits, they were matched for semantic similarity and thus redundancy within each musician, resulting in 215 statements (mean = 19.55, *SD* = 6.68). An overview of the amount of resulting thought-units and single statements before and after the redundancy check for each participant is given in Table 1.

Affinity diagramming. For the purpose of further reducing the pool of statements and to develop a first bottom-up categorical system of related statements, an affinity diagramming workshop was conducted by two HCI student research assistants from the same research lab. Affinity diagramming can be used for grouping information and generate categories in a bottom-up procedure (Beyer & Holtzblatt, 1999). In one four-hour workshop the 215 single statements were iteratively classified. In the beginning, the participants were briefed on how the data was generated in interviews and instructed to identify dimensions of the subjective quality of musical instruments. All the statements were placed on notes, making it easy to replace statements. The participants placed each statement at a time and dissolved disagreements by discussing the statement in more detail if necessary. During the matching of the statements, 18 were excluded because they did not match any of the emerging categories and were assumed to represent rather unimportant criteria. A pool of 197 statements remained at that stage. The four-hour workshop resulted in a first draft of the possible dimensions of the experiential quality criteria of musical instruments. The resulting categories and the number of statements of each category are shown in Table 2.

Further refinement and development of an initial item pool. It was decided to exclude the category ‘price’ (six statements) from the further development steps, since this category did not match the initial goal of the intended questionnaire. Based on the results of the grouped items, the refined categories were created by assigning each sub-level of the affinity diagram to a separate category (see Appendix A) This step made it easier to do a second check for semantic redundancy; this time within the produced categories itself and thus between the musicians who originally created those statements (before, we only looked for redundancy within the single musicians). Because the musicians often mentioned similar criteria, additional 64 statements were excluded from the pool of statements. The remaining 127 statements were used for generating an initial item pool. The item creation was achieved by following suggested guidelines for developing items by Thielsch, Lenzner and Melles (2012). The authors suggest specific rules for phrasing items for increasing their comprehensibility. The iteratively created items were consecutively assessed by two UX-experts from the same lab for correct language and phrasing issues which may produce problems in psychometric measurement. After that, the item pool included 100 items.

Table 1

Overview of interview participants.

Type of Musician	**Expertise**	**Plays music since (years)**	**Years of music education**	**Hours of music practice/ week**	**Thought-units**	**Single statements (total)**	**Single statements (reduced)**
Pianist	academic	22	17	25	34	43	32
Percussionist	academic	20	17	24	19	32	23
Pianist	semi-professional	18	10	6	22	23	15
Accordionist	semi-professional	30	7	4.5	12	14	7
Electronic musician	semi-professional	19	15	30	18	28	22
Saxophonist	semi-professional	25	13	8	18	18	13
Guitarist	recreational	20	10	1	33	38	20
Guitarist	recreational	15	9	3	32	37	22
Bassist	recreational	11	2	0.5	32	41	22
Electronic musician	recreational	18	7	5	22	23	15
Electronic musician	recreational	20	4	8	28	33	24
Sum		**218**	**111**	**115**	**270**	**330**	**215**
Mean		**19.82**	**10.09**	**10.45**	**24.55**	**30.00**	**19.55**
(Standard deviation)		(4.94)	(5.01)	(10.57)	(7.55)	(9.58)	(6.68)

Table 2

Solution of the three-level categorization (Affinity Diagramming Workshop).

Level 1	**Level 2**	**Level 3**	**Number of statements**
Hedonics			
	Aesthetics		11
	Mysticism		8
	Enjoyment		
		Fun	6
		Creativity	3
		Emotion/Expression	3
	Motivation		
		Challenge	11
		Freedom	21
		Surprise	9
Sound			7
	Timbre		8
	Quality		6
Playing experience			
	Playability		
		Personalization	4
		Control	12
		Hygiene	7
	Feel		
		Body/Ergonomics	11
		Body Elements	6
		Merging	8
	Usability		
		Agency	6
		Classic Usability	9
		Examination/Understanding	5
		Habit	7
Manufacturing			
	Material		6
	Tech. Manufacturing		12
	Tone/Quality		5
Price			6

Study 2: Refinement and reduction of the item pool

Study 2 aimed at further refining and reducing the item pool by asking experts to rate the importance of each item for assessing the subjective quality of musical instruments. The experts were also invited to suggest alternative phrasings and additional relevant items.

Method

Participants. The recruitment of experts took place via music research mailing lists and direct contacts through the provided research network. Forty-eight experts from the field of music research (mean age = 38.90, *SD* = 11.45, male = 40, female = 8) participated in an online study. The experts had various academic backgrounds: Design (N = 11), Music (N = 31), Academia, (N = 30), Engineering (N = 13), NIME (N = 22), and Other (N = 11)[1]. On a scale from 1 to 5 the experts rated their expertise on average with 4.00 (*SD* = 0.85)[2] On average the experts have been working in the field of music research for 13.34 years (*SD* = 11.52). One participant was excluded from the further analysis because of a low level of expertise (1: Fundamental Awareness). The remaining participants indicated at least a level of intermediate expertise and were thus included in the further analysis.

Materials. 100 items from the prepared item pool were subject of the analysis by the 47 experts. Experts had to rate each item regarding its importance on a 7-Point-Likert scale ranging from “Not at all important” to “Extremely important”.

Procedure. The music research experts were introduced to the study by explaining them that they were about to rate the importance of 100 statements (the items from the initial item pool) that would describe the subjective quality of musical instruments. We explicitly asked the experts to do this from their point of expertise. The experts could optionally suggest an alternative phrasing or leave a general comment on every item. After they rated the 100 items, we asked the experts if they would suggest further quality criteria for consideration. In the end, demographics were collected and a final question regarding

---

[1] Numbers do not add up to 48 since experts were allowed to indicate multiple academic backgrounds.
[2] The increments of the scale were 1 = Fundamental awareness, 2 = Novice, 3 = Intermediate, 4 = Advanced, and 5 = Expert.

the importance of a standardized questionnaire that assesses the subjective quality of musical instruments was posed.

Results Study 2

Because we intended to further refine the initial item pool with the overall mean importance ratings of the experts, we initially looked at the agreement among experts. Based on the level of agreement and the overall item importance we removed 24 items.

Interrater agreement. Overall interrater agreement was assessed by using Fleiss Kappa because of multiple raters. Experts only share a slight agreement (according to Landis & Koch, 1977)[3] on the importance of the items ($K = .045$, z-score = 11.75, $p < .001$) Because Fleiss Kappa is affected by the number of categories (Brenner & Kliebsch, 1996) the provided answer format of the 7-Point Likert scale was reduced to three mutually exclusive categories. Answer values from one to three were merged to the category “unimportant”, value four was labeled “neutral” and the remaining values from five to seven were grouped under the category “important”. Again Fleiss Kappa was calculated ($K = .119$, z-score = 54.67, $p < .001$) still representing only a slight overall agreement. Similar to that the assessment interrater reliability resulted in only a slight agreement with intra-class correlation (single measures): ($ICC(2,1) = .241$, $p < .001$). This finding is in alignment with a recent discussion about NIME-conference proceedings on evaluation (Barbosa et al., 2015) where researchers tend to have a different understanding of the term “evaluation”. Experts do not show an overall agreement of the importance of different criteria for the evaluation of the experiential quality of a musical instrument, therefore we looked at the proportion of agreement for each item in detail (Appendix B).

Proportion of agreement on item level. When using the three-category solution, experts show a higher absolute agreement for items that are rated more important. There is a strong correlation between the proportion of agreement of each item (p_i) and the average item importance (Pearson’s $r(98) = .92$, $p < .001$). In Figure 1, the relationship between proportion of agreement and item importance together with the included and excluded items

[3] Landis and Koch (1970) suggested interpretation of Kappa values that differentiate between six levels of agreement: poor (< .0), slight (.00 - .20), fair (.21 - .40), moderate (.41 - .60), substantial (.61 - .80) and almost perfect (.81 – 1.00).

is presented. The scatterplot shows that experts did agree less on items, which were overall rated as less important. On the contrary, items with a high overall importance indicate high interrater agreement.

Due to the low agreement on the seemingly less important items, the application of the mean importance rating as a decision rule for exclusion or retention of the items was not sufficient. We decided to consider a detailed approach to account for that difficulty. As shown in Table B (Appendix B), items were investigated with several exclusion criteria: (A) Relative majority of experts indicated a low importance, (B) the mean importance rating was below four and thus falling below the neutral middle category, (C) the item was unclear or problematic in terms of content (e.g. phrasing issues such as negative phrasing) and (D) in case of semantic redundancy items with a lower mean importance and higher proportion of agreement were removed. In the case of item Nr. 84 (Appendix A), the item was replaced with a slightly optimized one (Nr. 82, Appendix B). This procedure led to the exclusion of 24 items. Because of the suggestions from the experts four additional items were generated with regard to seemingly underrepresented categories. At the end, N = 80 items remained in the item pool.

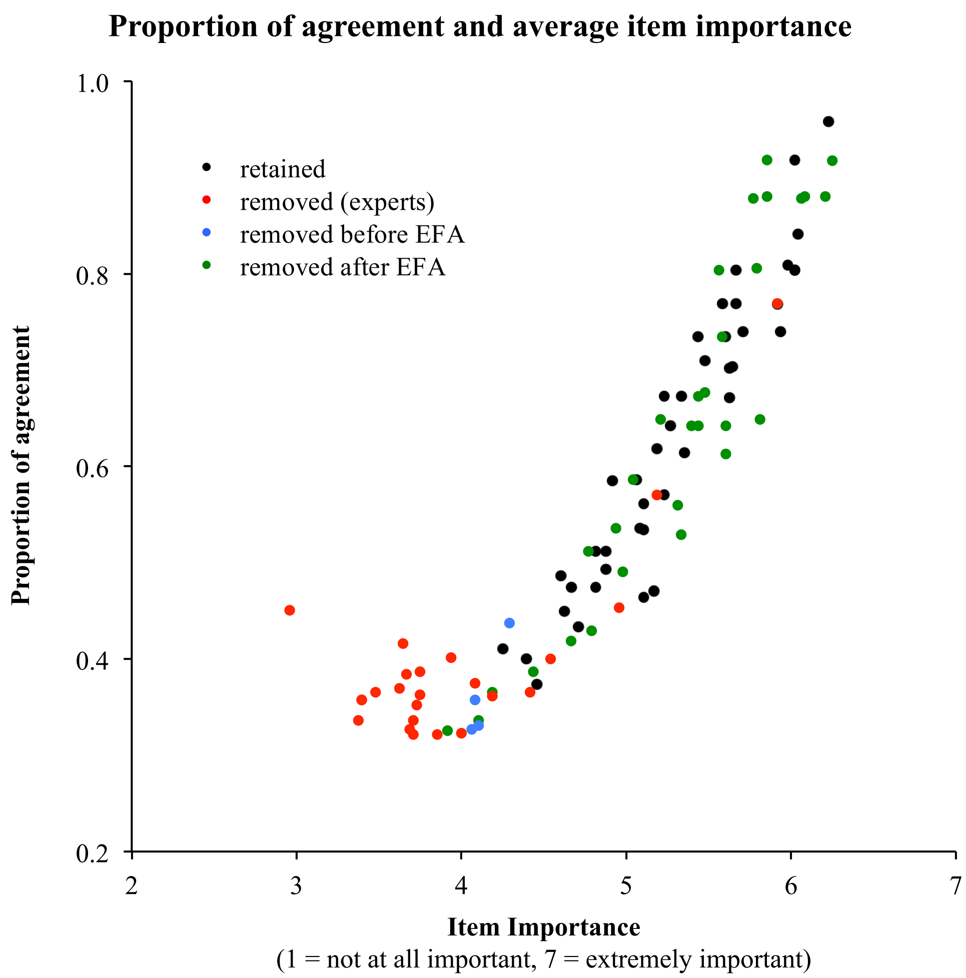

Figure 1. Scatterplot of the proportion of agreement and item importance. Note: The proportion of agreement is not corrected for agreement by chance and can thereby obtain a value between zero and one. Pearson's r = .92.

Study 3: Exploratory Factor Analysis

The final study of the present thesis aimed at identifying the underlying factor structure of the experienced musical instrument quality using exploratory factor analysis (EFA). With this approach correlative relations among variables are investigated, which allows for a reduction of the present items to fewer latent factors. Furthermore, EFA is often used for detecting inadequate items and provides decision rules for retaining or removing items.

Method

Participants. The recruitment of participants was performed in two ways. The first group (n = 133) was reached through an invitation letter distributed by e-mail, flyer and social-media posts. An e-mail campaign was run in order to reach professional musicians at music schools or similar institutions. The second group (n = 205) was recruited via a crowdsourcing platform. In order to enforce personal interest for participating in the survey, we decided to compensate the participants from the crowdsourcing platform with a rather low sum of money ($ 0.60). The data was cleaned using two attention checks (e.g. “please mark the strongly agree response”) and one honesty question which asked the participants if we should consider their data for our analysis. Both techniques are suggested for optimizing data quality in online surveys (Meade & Craig, 2012). Thirty-eight participants were excluded from the further analysis because they failed either in the attention check or indicated that we should not use their answers. Additionally, the data was screened for univariate, bivariate and multivariate outliers. After data cleaning, a sample of N = 300 participants remained for the exploratory factor analysis (180 male; 118 female; 2 without gender declaration). The participants were on average 33.19 years old (*SD* = 11.25) and indicated having 13.87 (*SD* = 11.88) years of experience on their instrument.

Materials. We prepared an online study where we included the refined set of the remaining 80 items (see bold-faced items in Appendix B). The items had to be rated on a 7-Point-Likert scale ranging from “Strongly disagree” to “Strongly agree”. We ordered the items alternatingly to prevent that items from theoretically similar facets were presented close together (e.g. items which related to control on a textual level where not placed after each other).

Procedure. After being introduced to the study, the participants were asked to remember a recent and memorable experience where they tested a musical instrument. The experience could explicitly be either good or bad. Since we are interested in correlations of similar items, we placed this advice with the intention to prevent from not generating enough variance in the data. In order to enhance the memory of the participants' experience, we asked them, where they were when they tested the instrument, how long ago their experience happened and for how long the experience lasted. This approach was inspired by Sheldon, Elliot, Kim and Kasser (2001) and Tuch, Trusell and Hornbæk (2013). After the musicians responded to all the questions we asked for various general information and musical demographics. In the end, the participants could leave questions or comments to the study. On average, the musicians spent 13.29 minutes (*SD* = 8.19 minutes) of their time for filling out the online survey.

Results study 3

The procedure towards the final factor solution was conducted with reference to a recently published guideline by Howard (2015) for realizing exploratory factor analysis (EFA) in the fields of cyberpsychology and human-computer interaction. Prior to the results from EFA, we present the experiences, which the factor solution is based on.

Participants' experiences. The class of the instruments the participants referred to in their experiences was assessed with two possible categorization systems. First, the predominant western taxonomy for categorizing musical instruments was (Sachs, 1965) and distinguishes musical instruments by the nature of their sound producing material. Second, a more common categorization of everyday language which differentiates the way musicians interact with the instrument was used. Most musicians remembered an experience where they used a chordophone instrument (169 indications) (Table 3). This accumulation is further reflected by the second category system where it is shown that keyboard, plucked and bowed and in some cases percussive instruments can account for that quantity. The values for all absolute frequencies and percentages for each category of those two taxonomies are given in Table 3. Furthermore, in Table 4 we present the frequencies for time and place of the memorable experiences. Most participants reported that the experiences happened some months ago, lasted between 15 to 30 minutes and most frequently took place either in their home or in a music store. Finally, we collected ratings (again on 7-Point Likert scales) for the overall impression of the instrument quality (ranging from (1) very bad to (7) very good; M = 5.85, *SD* = 1.22) and the experience (ranging from (1) very negative to (7) very positive; M = 5.82, *SD* = 1.31). Musicians mostly referred to instruments which they perceived as having a rather high quality which expectedly correlated significantly with their overall experience (Pearson's $r(298) = .633$, $p < .01$).

Table 3

Absolute frequencies and percent for the two classification systems of the instruments participants referred to in their experiences.

Type of classification	Sub-category	Absolute frequency	Percent (%)
Category of sound production			
	Aerophon[4]	60	20
	Chordophon[5]	169	56
	Electrophon[6]	45	15
	Idiophon[7]	12	4
	Membranophon[8]	14	5
	Total	**300**	**100**
Interaction category			
	Wind instrument	57	19
	Percussion instrument	29	10
	Bowed instrument	31	10
	Keyboard instrument	84	28
	Plucked instrument	99	33
	Total	**300**	**100**

[4] Aerophon: produces sound primarily by causing a body of air to vibrate (e.g. trumpet, flute, reed)

[5] Chordophon: sound is primarily produced by the vibration of strings (e.g. sitar, guitar, piano, violin)

[6] Electrophon: instruments involving electricity (e.g. theremin, hammond organ, synthesizer)

[7] Idiophon: set in motion by a percussive action: hitting, shaking, or scraping (e.g. marimba, vibraslap)

[8] Membranophon: instruments which have a struck membrane (e.g. conga, djembe, bodhrán)

Table 4

Absolute frequencies and percent on the three questions which intended to enhance the participants' experience about the specific experience they had.

Question	Absolute frequency	Percent
How long ago did you have this experience?		
Some hours ago	11	4
Some days ago	21	7
Some weeks ago	63	21
Some months ago	147	49
Some years ago	54	18
Within the last hour	4	1
Total	**300**	**100**
For how long did you use the instrument in that specific moment?		
less than 15 minutes	53	18
15-30 minutes	106	35
30-45 minutes	53	18
45-60 minutes	26	9
1-2 hours	37	12
2-4 hours	10	3
more than 4 hours	15	5
Total	**300**	**100**
Where were you when you had this experience?		
In my own home	85	28
In the home of friends	45	15
In a music store	74	25
At an instrument designer	15	5
At a concert	10	3
At a rehearsal	29	10
At an instrument exhibition	3	1
In a music lesson	20	7
Other	19	6
Total	**300**	**100**

Item Analysis. The items were inspected for inter-item correlations, discriminatory power and item difficulty. This step allows for identifying inappropriate items before running EFA (Moshagen & Thielsch, 2010). Tabachnick and Fidell (2007) suggested to scan the item correlation matrix for a reasonable amount of correlations of at least above .30. Although the average inter-item correlations tended to be in an acceptable range, we applied a more conservative benchmark. Hence, it was decided to exclude items with average inter-item correlations below .40 which resulted in removing five items from the data set (ranging from .29 to .39). In conformance to that analysis, these five items had the lowest discriminatory power (corrected item-total correlation: $r < .50$). The item difficulties were all in an acceptable range of .15 to .85 (e.g. Moshagen & Thielsch, 2010). Subsequently, 75 items remained in the item pool.

EFA Assumptions. EFA requires adequate sample sizes which is an often-discussed issue among psychometricians (e.g. MacCallum, Widaman, Zhang & Hong (2001) According to Comrey and Lee (1992) a sample size of 100 is poor, 200 is fair, 300 is good, 500 is very good and 1000 or above is excellent. Furthermore, it is suggested to have a subject to item ratio of at least 3:1, but more reliable solutions are derived from bigger ratios. With the remaining number of items and participants, we reached a subject to item ratio of 4:1 ($N = 300$, number of items = 75). Although this ratio is at the lower end of the recommended sizes, MacCallum, et al. (2001) argued that if the communalities are high, the importance of this requirement decreases. The communalities in our case were middle to high, ranging from .357 to .764 ($M = .61$, $SD = .10$). Prior to the analysis we examined the factorability of the 75 items. All items correlated at least .30 with each other. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy reached a value of .97 over all items, above the suggested value of .60 for all 75 items and Bartlett's test of Sphericity was significant ($\chi2(2775) = 23398$, $p < .001$). Finally, the anti-image correlation diagonals were all above .90.

Factor Retention. Generally, it is recommended to use more than one criterion for guidance in the decision of factor extraction (Howard, 2015). First, researchers are advised against the usage of the Kaiser-Criterion where factors with an eigenvalue greater than one are extracted, because following this rule would often lead to overestimating the number of factors (Howard, 2015). Besides the investigation of the Scree-Plot, Howard mentions that Parallel Analysis provides more reliable solutions. The eigenvalues from Parallel Analysis can be used to determine if the real eigenvalues from the researcher's dataset are beyond chance. Until the point where the real eigenvalue exceeds the eigenvalue from the Parallel Analysis, factors are extracted (O'Connor, 2000). The results from that suggested a four-factor solution which was in accordance to the Scree-Plot. To account for the possibility of over- or underestimating the number of factors, solutions from three to six factors were compared. The interpretability and cleanness of the solutions increased to the low end when less factors where extracted and an emerging three-factor solution was evident. Because the items are all expected to measure the same overall dimension, principal axis factoring (PAF) with oblique rotation (direct oblimin) was calculated allowing the factors to correlate with each other. All analyses were computed with IBM SPSS 23.

Factor loading Cutoff. Howard (2015) suggested to retain items that show: primary factor loadings > .40, alternative loadings below .30 and a greater difference than .20 between the primary and alternative factor loadings. Following this guideline, the three-factor solution was incrementally refined (each item at a time). Because one factor showed signs for overrepresentation (Eigenvalue of 40.7), a more conservative adaption of Howard's recommendations was applied. The primary loadings had to exceed a value of .50 to be retained in the solution. For the final solution 32 items were removed. Hence, the presented factor structure in Table 5 contains the remaining 43 items. The factors show as expected substantial inter-correlations (all below .7) (Table 6).

Table 5

Pattern matrix of the final factor solution with the remaining 43 items after deleting items with primary loadings < .5, alternative loadings > .3 , delta of loadings < .2 between the primary and alternative loadings and communalities.

Category[a]	**Item**	**Factor Loading Pattern**			***h^2***
Factor 1: Experienced freedom and possibilities		F1	F2	F3	
Freedom	The instrument allows me to learn new things	**0.88**			0.53
Freedom	The instrument offers me new possibilities of things to do	**0.85**			0.45
Explorability	I can continually discover new things by using the instrument	**0.83**			0.63
Freedom	The instruments expands my experience of musical interaction	**0.82**			0.70
Freedom	The instrument offers me new facets of playing	**0.81**			0.49
Creativity	The instrument fosters my creativity	**0.73**			0.63
Freedom	I perceive the instrument as offering a lot of variety	**0.72**			0.65
Expressiveness	The instrument offers new possibilities to express myself musically	**0.71**			0.64
Challenge	I perceive the instrument as challenging in a positive way	**0.70**			0.62
Freedom	The instrument feels like I can go beyond myself	**0.69**			0.65
Freedom	The instrument offers me interesting possibilities to manipulate sound	**0.66**			0.70
Challenge	Playing the instrument allows me to further develop my musical skills	**0.63**			0.46
Engagement	The instrument keeps me interested	**0.62**			0.66
Engagement	The instrument allows me to be engaged when I am playing it	**0.59**			0.75
Freedom	The instrument offers me a good variety of sounds to evoke	**0.58**			0.57
Comfort	The instrument allows me to focus on sound generation	**0.54**			0.61
Creativity	The instrument supports me in creating new music in any style	**0.52**			0.67
Factor 2: Perceived control and comfort					
Control	I feel in control of the instrument		**0.95**		0.52
Control	I can play precisely on the instrument		**0.81**		0.69
Conformance	The instrument does what I want it to do		**0.79**		0.49
Control	I can control the sound appropriately		**0.78**		0.51
Conformance	I can use the instrument intuitively		**0.68**		0.44
Play Comfort	I feel relaxed when I play the instrument		**0.66**		0.53
Ergonomics	I feel comfortable when I play the instrument		**0.65**		0.68
Conformance	The instrument responds well to my actions		**0.64**		0.52
Play Comfort	I perceive the instrument as comfortable to play		**0.63**		0.57
Play Comfort	The material allows me to use comfortable gestures		**0.60**		0.65
Conformance	I feel like I am initiating, executing and controlling the behavior of the instrument		**0.57**		0.60
Ergonomics	I perceive the instrument as allowing small and efficient movements		**0.57**		0.67
Conformance	The instrument works the way I expect it to		**0.57**		0.75
Ergonomics	The instrument feels like an extension to my body		**0.57**		0.71
Engagement	It is easy for me to get into a flow of playing with the instrument		**0.56**		0.75
Factor 3: Perceived stability, sound quality and aesthetics					
Stability	I think the instrument is very well manufactured			**0.81**	0.69
Sound Quality	The instrument produces high quality sound			**0.79**	0.68
Stability	I think the instrument is reliable			**0.74**	0.69
Stability	I perceive the instrument as solid			**0.65**	0.53
Sound Quality	The sound quality is convincing to me			**0.64**	0.70
Aesthetics	The instrument looks appealing			**0.62**	0.73
Sound Quality	The instrument pleases me sound-wise			**0.60**	0.71
Aesthetics	I find the instrument aesthetically pleasing			**0.58**	0.69
Sustainability	I perceive the instrument as sustainable in a musical way			**0.58**	0.68
Stability	I think the instrument is able to endure a lot			**0.57**	0.55
Sound Quality	I think the material of the instrument positively supports the sound generation			**0.56**	0.61

Notes. Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. a Rotation converged in 9 iterations. Note: only loadings > .3 are shown. [a] Category refers to the refined categories of study 2.

Table 6

Factor correlation matrix of the final solution.

Factor Correlation Matrix

Factor	EFP	PCC	PSSQA
EFP	1.000	.599	.657
PCC	.599	1.000	.645
PSSQA	.657	.645	1.000

Table 7

Descriptive statistics for the three subscales.

Factor / Subscale	Number of items	Scale M *(SD)*	Eigenvalue λ	Percent of explained Variance	Cronbach's α
EFP	17	91.61 (18.71)	22.80	53.03	0.97
PCC	15	80.11 (16.41)	3.10	7.21	0.96
PSSQA	11	62.50 (11.36)	1.95	4.53	0.94

Semantic content of the resulting three-factor solution. The evolved factor structure from the EFA suggests three interrelated facets. (1) The first and statistically predominant factor (accounting for a majority of the explained variance, see Table 7) contains 17 items. Regarding the refined categories from study 1, eight items are related to the aspect of freedom, one item accounts for explorability, two for creativity, one for expressiveness, further two for each challenge and engagement, and a final one for comfort. Except for the single comfort item, these items were all initially placed in the affinity diagramming groups of motivation and enjoyment, which we initially saw as two sub-groups of musical instrument hedonics. (2) The second factor, which holds 15 items includes three items of control, five items regarding conformance, three items of play comfort, three items that account for ergonomics and one item that was labeled as engagement. These items were, except for the engagement item, derived from the statements grouped under playability and feel as sub-levels of the playing experience.
(3) The third factor with eleven items accounts for stability with three items, with four items for sound quality and two items describing aesthetic aspects. In contrast to the other two factors, these items were derived from three different 1. level affinity diagram groups (Manufacturing, Sound and Aesthetics).

General Discussion

Out of 100 items, which were derived from 330 single statements of musicians in study one, we developed a questionnaire containing 43 items measuring three interrelated facets of the musician's perception of the experiential quality of musical instruments (MPX-Q). The subscales show high internal consistencies, indicating a reliable measurement tool for the psychometric assessment of the experiential quality of different musical instruments. The present findings suggest that three interrelated latent facets contribute to this higher-level construct. (1) *Experienced freedom and possibilities* (EFP) as a predominant facet relates to exploring new things in music through the help of a musical instrument. The corresponding items converge to an underlying motive of finding new ways to develop musicianship and musical expressivity. (2) *Perceived control and comfort* (PCC) describes musical controllability which is an often assessed facet of interest in the previous literature (e.g. Wanderley & Orio, 2002). PCC also refers to ergonomic aspects which contribute to the perceived comfort of a musical instrument. (3) *Perceived stability, sound quality and aesthetics* (PSSQA) is the most heterogeneous facet of the resulting three. It seems to include rather classic instrument quality notions that are found in the material, the sound or the visual appearance of an instrument.

With reference to our research goals, we explored the underlying structure of items that evolved in a bottom-up procedure from interviews with musicians. By letting music research experts evaluate the importance of those items we were able to identify obsolete ones. We were able to work towards the finding of the key experiential qualities of musical instruments from the musician's perspective. This was done running an online survey with various musicians, rating a memorable experience with a musical instrument.

Interestingly, the found factor structure shows a notable similarity to the conceptualization of violin quality by violinists (Saitis, 2014). Saitis conducted a lexical analysis of interview data from studies that investigated perceptual evaluation of violins (Saitis, Giordano, Fritz, & Scavone, 2012; Saitis, Scavone, Fritz, & Giordano, 2015). Four dimensions of violin quality were detected: *Handling*, *Sound*, *Relevance* and *Balance Across Strings*. Because the latter category appears to be a rather specific category in violin quality conceptualization, we discuss the first three categories in relation to the results of our analysis. Furthermore, we look at the relevance of these facets in the context of the related work of NIME-research.

Experienced freedom and possibilities (EFP). A majority of the items in this facet corresponds to the notion of developing new skills by enabling the musician to explore new and unknown possibilities and features of a musical instrument. Although we labeled only one of the items with 'expressiveness', this concept also matches the content of the other EFP-items because expressiveness seems to suit an underlying motive of this facet. In our view this facets resembles the dimension of *Relevance* introduced by Saitis (2014). According to Saitis, relevance consists of two aspects: first, *affective reactions* which relate to hedonic aspects such as interestingness, enjoyment, fascination or beauty, and second, *musical and emotive potential*, including characteristics of expressivity that the instruments allows the musician to convey. With reference to Justin's (2003) GERMS model that focuses more on performative aspects of expressivity, the content of the EFP-items rather corresponds to the musician's impression of the instrument to hold new and desired possibilities that foster the musical development of the musician (e.g. generating and applying new sounds or varying the musical playing as desired). From the point of view that expressivity relies on having diverse musical possibilities, the common notion of the EFP-items therefore refers also to the perceived freedom of choice between different available musical possibilities and their novelty to the individual musician.

Perceived control and comfort (PCC). PCC consists almost exclusively of attributes related to control and comfort by attaching weight to ergonomic aspects of musical instruments. Musicians tend to highly associate the aspects of control with ergonomics and conformance. An instrument with a high controllability allows to evoke the intended sounds and thereby is in conformance to the expectation of the performer. Therefore, one could also refer to it as "musical agency". The PCC facet found in this study shows a high similarity to the *Handling* dimension identified by Saitis (2014); both are related to the overall playability of a musical instrument. Handling, includes *response* and *design & comfort.* Response includes aspect like *ease of playing* and *reaction* or *well articulation*. For the violin this is of particular relevance since a number of physical factors contribute to bow quality (Matsunaga, Sugiyama, Minato, & Norimoto, 1996) as well as playing parameters (O'Modhrain, Serafin, Chafe, & Smith, 2000). In contradistinction to our findings, Saitis (2014) also lists *liberty* among those aspects. Our results suggest that this aspect is also relevant under the consideration of the facet EFP as a rather hedonic attribute of musical instrument quality in distinction to control and ergonomics. Because of the high interrelation of the facets it is however also possible that control and ergonomics serve as requirement for perceived liberty. The second aspect of Handling refers to the feel of the instrument, which is determined by design aspects, such as size, shape, weight or comfort. In NIME-research, the concept of feel was aimed to be operationalized by Marshall and Wanderley (2011) for evaluating a haptic augmented instrument. Marshall and Wanderley assessed ease of use, controllability, engagement, entertainment and potential for further performance. Although the authors provide no specific reasoning for this choice of characteristics, controllability and ease of use would match the PCC facet, whereas engagement, entertainment and potential rather fit the EFP facet due to their hedonic nature. The increasing interest in the investigation of the role that the haptic modality occupies in musical interaction (Järveläinen, Papetti, Schiesser, & Grosshauser, 2013) is of particular relevance for this facet for instrument evaluation. It is mentioned that instrumentalists use vibrotactile cues for enhanced control and performance (D. M. Birnbaum & Wanderley, 2007). Therefore we assume, that the PCC facet may play an important role in haptic music research and should be further investigated for sensitivity for changes in haptic feedback.

Perceived stability, sound quality and aesthetics (PSSQA). Concerning the initially assumed refined categories which resulted at the end of study one (manufacturing, sound quality and aesthetics) this facet appears to be the most heterogeneous in comparison to the two previously discussed facets. In the initial affinity diagram, the category aesthetics was grouped under hedonics, sound quality under sound and stability under manufacturing. With reference to Saitis' three-dimension space only items regarding sound quality overlap with Saitis' sound dimension which itself accounts much more detailed for different characteristics of violin sound than the present items on sound quality do. However, the high interrelations of sound quality, instrument stability and aesthetics may point towards another interesting phenomenon. Although our results are only correlational, it may be valuable to investigate the relationships between perceived visual or haptic instrument aesthetics (the musical instrument's look and feel) and its potential influence on the perceived sound quality, stability or the other interrelated facets of the MPX-Q. In HCI, the claim of an aesthetic bias on usability (Tractinsky, Katz, & Ikar, 2000) based on a "halo-effect" which stems from social psychology (Dion, Berscheid, & Walster, 1972) led to a controversial discussion. The effect describes that physical beauty is obvious and highly accessible and thus can be transferred to personality traits, which can lead to a biased impression about a person (e.g. biased trustworthiness). In HCI, however, it was also shown that the effect can be reversed by usability influencing the aesthetic impression (Tuch, Roth, Hornbæk, Opwis, & Bargas-Avila, 2012). With some similarity, in the domain of traditional musical instrument research a preference bias was suggested to be influenced by monetary value and age (Fritz, 2012) because controlling for this effect by letting musicians play violins blindfolded led participants to prefer new violins over old Italians. Therefore it seems probable that further aspects, beyond monetary value and age could contribute to a biased impression of musical instrument quality. Visual aesthetics may be one them because 10 out of 11 interviewed musicians mentioned that a musical instrument would had to look good. One participant added that aspect of visual aesthetics would not play a central role but would often be a hint for a good quality instrument.

Summarized, we be believe that the MPX-Q questionnaire will help designers and researchers in the field of NIME to demonstrate how musicians perceive musical instruments on the three quality facets in a more profound way than with the application of ad-hoc scales. The questionnaire can be helpful for the investigation of low to high level prototypes and finished instruments and contribute to the guidance in design decisions. Designers could address if their instruments correspond to their initial intentions and researchers can use the questionnaire for experimental studies where the overall quality of musical instruments is an outcome of interest. Furthermore, we believe that the three quality facets can contribute to a nuanced understanding of musical instrument quality in general, because they are expected to be universal to some extent, similar to psychometric tools in the field of HCI.

Suitability of the questionnaire for NIME-research

Our results are based on data from different, but mostly very common musical instruments. The reasons behind that were twofold. First, due to practical reasons we aimed at surveying musicians with various backgrounds and therefore it was not exclusively required to report an experience with regard to a NIME. Achieving a similar amount of participants only with NIMEs requires more resources because of their low availability. The second reason was that we were inspired by the fact that in HCI-research psychometrics are applicable over a broad range of interactive products (Hassenzahl et al., 2003), which led us to the assumption of searching for key quality criteria of musical instruments that could potentially be universal to some extent. Although the classification of musical instruments in different sub-groups has a long extensive history (Sachs, 1965) we believe that the claim for an general-purpos quality comprehension is neither bound to distinct categories of musical instruments nor the various musical styles and genres. This assumption is supported by comparable approaches in HCI and the continuous endeavor for valid and reliable measurement of different cognitive constructs in Psychology (e.g. Messick, 1995) that have proven to be stable conceptualizations of the human mind (e.g. universal judgments of emotions based on facial expressions; Ekman et al., 1987). One could argue that this relation is far-fetched and thus not true for the domain of musical instrument evaluation and NIMEs in particular. However, similar to Dobrian and Koppelman's (2006) suggestion for a NIME-design approach that is inspired by existing acoustic instruments, the development of our questionnaire was shaped in a corresponding way by the conceptualization of musical instrument quality of existing instruments.

Looking at the DMI classification system of Miranda and Wanderley (2006) it seems adequate to assume a higher suitability of MPX-Q for instruments with a greater degree of instrument similarity (ranging from instrument-like controllers to alternate controllers). This is supported by Young and Murphy (2015) who explain that designers actively avoid implementing familiarity to existing instruments in new alternate controllers, whereas in the first three categories this notion is loosed. Beyond that, it has been argued that instruments from the NIME-community are often designed to meet idiosyncratic needs of the performers or composers (Orio & Wanderley, 2002). This often leads to the development of one-of-a-kind devices (Papetti, Järveläinen, & Grosshauser, 2013) where the application of psychometric evaluation may not be suitable. Hence, we assume that the MPX-Q questionnaire is more suitable for a holistic evaluation of whole DMIs that are intended to be reflected by numerous musicians and provide some sort of familarity to already exisiting musical instruments.

Limitations and future research

The presented findings are subject to several limitations. First, although we aimed at being as inclusive as possible in finding quality criteria that are relevant for the perceived experiential quality of musical instruments, there may be further important aspects which were not implemented in the questionnaire development process. Second, the present questionnaire needs to be subject of further refinement. Several items still seem to overlap on a textual level. This version of the questionnaire is still quite long and could be further refined and shortened. A more concise questionnaire prevents from causing fatigue and can provide more reliable data (Galesic & Bosnjak, 2009) Third, although we found a three-factor structure that accounts for different musical instruments, it has yet to be proven that this structure shows consistency over different groups of musical instruments by conducting Confirmatory Factor Analysis for assessing construct validity (Harrington, 2008). Furthermore, the construct validity of the questionnaire needs to be investigated in terms of convergent and divergent validity - with similar and different constructs by using psychometric and psychophysiological measurements. Additionally, discriminant validity which accounts for the ability of the psychometric tool to effectively differentiate between different musical instruments (e.g. at different design stages) needs to be investigated in future research (Nunnally & Bernstein, 1994). Subsequently, we plan to use the questionnaire for experimental investigations of the effect of haptic feedback on the experiential quality of haptically augemnted DMIs (Papetti et al., 2015). Lastly, we want to emphasize that this work is only a first step towards the demanding objective of measuring musical instrument quality from the musician's perspective.

Conclusion

With regard to the current lack of available psychometric measurements in the field of NIME-research we provide an initial questionnaire for the assessment of the musician's perceived experiential qualities of musical instruments (MPX-Q). Based on the collection of a broad range of possible quality criteria, we carefully reduced our initial item pool to identify the key quality criteria by asking experts to rate the importance of each item. The final solution which is based on memorable experiences with various musicians suggest three interrelated facets of MPX-Q: Experienced freedom and possibilities (EFP), Perceived control and comfort (PCC) and Perceived stability, sound quality and aesthetics (PSSQA). The three scales of the three facets show high internal constancies. Further investigations are needed that assess the validity and the practicability of the suggested questionnaire.

References

Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Waveland Press.

Barbosa, J., Malloch, J., Huot, S., & Wanderley, M. M. (2015). What does "evaluation" mean for the nime community? In *New Interfaces for Musical Expression (NIME)*.

Beyer, H., & Holtzblatt, K. (1999). Contextual design. *Interactions*, *6*(1), 32–42. doi: 10.1145/291224.291229

Birnbaum, D., Fiebrink, R., Malloch, J., & Wanderley, M. M. (2005). Towards a dimension space for musical devices. In S. S. Fels, T. Blaine, A. Schloss, & S. Jordà (Eds.), *New Interfaces for Musical Expression (NIME)* (pp. 192–195). Vancouver, BC, Canada.

Birnbaum, D. M., & Wanderley, M. M. (2007). A systematic approach to musical vibrotactile feedback. In *Proc. Int. Computer Music Conf. (ICMC)*.

Bongers, B. (2000). Physical Interfaces in the electronic arts: Interaction theory and interfacing techniques for for real-time performance. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 41–70). Paris: Ircam - Centre Pompidou.

Brenner, H., & Kliebsch, U. (1996). Dependence of weighted kappa coefficients on the number of categories. *Epidemiology*, *7*(2), 199–202.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285.

Ekman, P., Friesen, W. V, O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*. US: American Psychological Association. doi: 10.1037/0022-3514.53.4.712

Erkut, C., Jylhä, A., & Discioglu, R. (2011). A structured design and evaluation model with application to rhythmic interaction displays. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.

Franinović, K., & Serafin, S. (2013). *Sonic interaction design*. Mit Press.

Fritz, C., Curtin, J., Poitevineau, J., Borsarello, H., Wollman, I., Tao, F.-C., & Ghasarossian, T. (2014). Soloist evaluations of six old italian and six new violins. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(20), 7224–9. doi: 10.1073/pnas.1323367111

Fritz, C., Curtin, J., Poitevineau, J., Morrel-Samuels, P., & Tao, F.-C. (2012). Player preferences among new and old violins. *Proceedings of the National Academy of*

Sciences of the United States of America, *109*(3), 760–3. doi: 10.1073/pnas.1114999109

Fritz, C., & Dubois, D. (2015). Perceptual evaluation of musical instruments: State of the art and methodology. *Acta Acustica United with Acustica*, *101*(2), 369–381. doi: 10.3813/AAA.918833

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. doi: 10.1093/poq/nfp031

Greenberg, S., & Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (p. 111). New York, New York, USA: ACM Press. doi: 10.1145/1357054.1357074

Harrington, D. (2008). *Confirmatory factor analysis*. Oxford University Press, USA.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. In *Mensch & Computer 2003* (pp. 187–196). Vieweg + Teubner Verlag.

Hassenzahl, M., Kekez, R., & Burmester, M. (2002). The importance of a software's pragmatic quality depends on usage modes. In *Proceedings of the 6th international conference on Work With Display Units (WWDU 2002)* (pp. 275–276).

Howard, M. (2015). A review of exploratory factor analysis (efa) decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction*, doi: 10.1080/10447318.2015.1087664

Järveläinen, H., Papetti, S., Schiesser, S., & Grosshauser, T. (2013). Audio-Tactile feedback in musical gesture primitives: Finger pressing. In *Sound and Music Computing (SMC)* (pp. 109–114).

Johnston, A., Candy, L., & Edmonds, E. (2008). Designing and evaluating virtual musical instruments: facilitating conversational user interaction. *Design Studies*, *29*(6), 556–571. doi: 10.1016/j.destud.2008.07.005

Jordà, S. (2004). Digital instruments and players: part I - efficiency and apprenticeship. In *New Interfaces for Musical Expression (NIME)* (pp. 59–63).

Jordà, S., & Mealla, S. (2014). A methodological framework for teaching, evaluating and informing nime design with a focus on mapping and expressiveness. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 233–238). London, United Kingdom: Goldsmiths, University of London.

Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music*, *31*(3), 273–302. doi: 10.1177/03057356030313003

Kiefer, C., Collins, N., & Fitzpatrick, G. (2008). HCI methodology for evaluating musical controllers: A case study. In A. Camurri, G. Volpe, & S. Serafin (Eds.), *Proceedings of the 2008 International Conference on New Interfaces for Musical Expression (NIME-08)* (pp. 87–90). Genoa, Italy.

Kontogeorgakopoulos, A., & Kouroupetroglou, G. (2011). Simple cases of low cost force-feedback interaction with haptic digital audio effects. In *International Gesture Workshop* (pp. 142–146). Athens.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. http://doi.org/10.2307/2529310

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis.

Malloch, J., Birnbaum, D., Sinyor, E., & Wanderley, M. M. (2006). Towards a new conceptual framework for digital musical instruments. In *Proceedings of the 9th Int. Conference on Digital Audio Effects* (pp. 49–52). Montreal, Canada.

Marshall, M. T., & Wanderley, M. M. (2011). Examining the effects of embedded vibrotactile feedback on the feel of a digital musical instrument. In *New Interfaces for Musical Expression* (pp. 399–404).

Matsunaga, M., Sugiyama, M., Minato, K., & Norimoto, M. (1996). Physical and mechanical properties required for violin bow materials. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, *50*(6), 511–517.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*. Meade, Adam W.: Department of Psychology, North Carolina State University, Campus Box 7650, Raleigh, NC, US, 27695-7650, American Psychological Association. doi: 10.1037/a0028085

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9).

Miranda, E. R., & Wanderley, M. M. (2006). *New digital musical instruments: Control and interaction beyond the keyboard*. Midelton, Wisconsin: AR Editions.

Morreale, F., Angeli, A. De, & O'Modhrain, S. (2014). Musical interface design: An experience-oriented framework. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 467–472). London, United Kingdom: Goldsmiths, University of London.

Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, *68*(10), 689–709. doi: 10.1016/j.ijhcs.2010.05.006

Nunnally, C., & Bernstein, H. (1994). Psychometric theory. New York: McGraw-Hill.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *32*(3), 396–402.

O'Modhrain, S. (2011). A framework for the evaluation of digital musical instruments. *Computer Music Journal*, *35*(1), 28–42.

O'Modhrain, S., Serafin, S., Chafe, C., & Smith, J. O. (2000). Influence of attack parameters on the playability of a virtual bowed string instrument: tuning the model. In *International Computer Music Conference (ICMC)*.

Orio, N., & Wanderley, M. M. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*. doi: 10.1162/014892602320582981

Overholt, D. (2009). The musical interface technology design space. *Organised Sound*, *14*(02), 217. doi: 10.1017/S1355771809000326

Papetti, S., Järveläinen, H., & Grosshauser, T. (2013). Effects of audio-tactile feedback on force accuracy in a finger pressing task: a pilot study. In *Haptic and Audio Interaction Design (HAID)*.

Papetti, S., Schiesser, S., & Fröhlich, M. (2015). Multi-point vibrotactile feedback for an expressive musical interface. In *New Interfaces for Musical Expression (NIME)*.

Poepel, C. (2005). On interface expressivity: a player-based study. In *New Interfaces for Musical Expression (NIME)* (pp. 228–231). Vancouver, BC, Canada: National University of Singapore.

Poupyrev, I., Lyons, M. J., Fels, S., & Blaine (Bean), T. (2001). New interfaces for musical expression. In *CHI '01 extended abstracts on Human factors in computing systems - CHI'01* (p. 491). New York, New York, USA: ACM Press. doi: 10.1145/634067.634348

Reynolds, T. J., & Gutman, J. (1988). Laddering theory, method, analysis, and interpretation. *Journal of Advertising Research*, *28*(1), 11–31.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1).

Sachs, C. (1965). *The history of musical instruments*. New York: Courier Dover Publications.

Saitis, C. (2014). *Evaluating violin quality: player reliability and verbalization*. McGill.

Saitis, C., Giordano, B. L., Fritz, C., & Scavone, G. P. (2012). Perceptual evaluation of violins: a quantitative analysis of preference judgments by experienced players. *The Journal of the Acoustical Society of America*, *132*(6), 4002–12.

http://doi.org/10.1121/1.4765081

Saitis, C., Scavone, G. P., Fritz, C., & Giordano, B. L. (2015). Effect of task constraints on the perceptual evaluation of violins. *Acta Acustica United with Acustica*, *101*(2), 382–393. doi: 10.3813/AAA.918834

Schmid, G.-M. (2014). Measuring musician's playing experience: Development of a questionnaire for the evaluation of musical interaction. In *Practice-Based Research Workshop at NIME*. London, UK.

Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology*, *80*(2), 325.

Stowell, D., Plumbley, M. D., & Bryan-Kinns, N. (2008). discourse analysis evaluation method for expressive musical interfaces. In A. Camurri, G. Volpe, & S. Serafin (Eds.), *International Conference on New Interfaces for Musical Expression (NIME)* (pp. 81–86). Genoa, Italy.

Stowell, D., Robertson, A., Bryan-Kinns, N., & Plumbley, M. D. (2009). Evaluation of live human–computer music-making - Quantitative and qualitative approaches. *International Journal of Human-Computer Studies*, *67*(11), 960–975. doi: 10.1016/j.ijhcs.2009.05.007

Swift, B. (2013). Chasing a feeling: Experience in computer supported jamming. In *Music and human-computer interaction* (pp. 85–99). Springer London.

Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics. Boston: Pearson Education

Thielsch, M. T., Lenzner, T., & Melles, T. (2012). Wie gestalte ich gute items und interviewfragen? *Praxis Der Wirtschaftspsychologie II. Themen Und Fallbeispiele Für Studium Und Anwendung. Monsenstein Und Vannerdat, Münster*, 221–240.

Tractinsky, N., Katz, A. ., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13*(2), 127–145. doi: 10.1016/S0953-5438(00)00031-X

Tuch, A. N., Roth, S. P., Hornbæk, K., Opwis, K., & Bargas-Avila, J. A. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior*, *28*(5), 1596–1607. doi: 10.1016/j.chb.2012.03.024

Tuch, A. N., Trusell, R., & Hornbæk, K. (2013). Analyzing users' narratives to understand experience with interactive products. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (p. 2079). New York, New York, USA: ACM Press. doi: 10.1145/2470654.2481285

Wallis, I., Ingalls, T., Campana, E., & Vuong, C. (2013). *Amateur Musicians, Long-Term Engagement, and HCI. Music and Human-Computer Interaction*. doi: 10.1007/978-1-

4471-2990-5_3

Wanderley, M. M., & Orio, N. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, *26*(3), 62–76. doi: 10.1162/014892602320582981

Wollman, I., Fritz, C., & Poitevineau, J. (2014). Influence of vibrotactile feedback on some perceptual features of violins. *The Journal of the Acoustical Society of America*, *136*(2), 910. doi: 10.1121/1.4889865

Young, G., & Murphy, D. (2015). HCI models for digital musical instruments: methodologies for rigorous testing of digital musical instruments. In *International Symposium on Computer Music Multidisciplinary Research (CMMR)*. Plymouth.

Appendix A

Table A

Sorted Items and their respective categories from study 1: the final refined categories and the categories from the affinity diagram divided in Level 1, 2 & 3 (a.d.).

Nr.	Item Text	Refined categories	Level 1 (a.d.)	Level 2 (a.d.)	Level 3 (a.d.)
1	I find the instrument aesthetically pleasing	Aesthetic Impression	Hedonics	Aesthetics	
2	I perceive the instrument as flawless	Aesthetic Impression	Hedonics	Aesthetics	
3	The instrument looks appealing	Aesthetic Impression	Hedonics	Aesthetics	
4	I perceive the instrument as unique	Aesthetic Impression	Hedonics	Aesthetics	
5	The colors of the instrument are appealing	Aesthetic Impression	Hedonics	Aesthetics	
6	The instrument allows me to be creative	Creativity	Hedonics	Fun	Creativity
7	The instrument supports me in creating new music	Creativity	Hedonics	Fun	Creativity
8	The instrument fosters my creativity	Creativity	Hedonics	Fun	Creativity
9	The instrument supports me in releasing my inner muse	Creativity	Hedonics	Fun	Creativity
10	I perceive the instrument as easy to use	Ease of Use	Playing experience	Usability	Classic Usability
11	I can personalize the instrument	Ease of Use	Playing experience	Playability	Personalization
12	I perceive the instrument as unnecessarily complex	Ease of Use	Playing experience	Usability	Classic Usability
13	The instrument allows me to be engaged when I'm playing it	Engagement	Playing experience	Feel	Merging
14	I can imagine to play the instrument for a long time	Engagement	Motivation	Challenge	
15	I feel the urge to play the instrument again	Engagement	Motivation		
16	The instrument keeps me interested	Engagement	Motivation	Surprise	
17	It is easy for me to get into a flow of playing with the instrument	Engagement	Motivation	Challenge	
18	The instrument surprises me in a positive way	Engagement	Motivation	Surprise	
19	I have fun playing the instrument	Enjoyment	Hedonics	Fun	

Nr.	Item Text	Refined categories	Level 1 (a.d.)	Level 2 (a.d.)	Level 3 (a.d.)
20	The instrument responds well to my actions	Conformance	Playing experience	Feel	Self-induced causality
21	I feel like I am initiating, executing and controlling the behavior of the instrument	Conformance	Playing experience	Feel	Self-induced causality
22	The instrument does what I want it to do	Conformance	Playing experience	Feel	Self-induced causality
23	The instrument works the way I expect it to	Conformance	Playing experience	Feel	Self-induced causality
24	I can use the instrument intuitively	Conformance	Playing experience	Usability	Classic Usability
25	The instrument allows me to express myself	Expressiveness	Hedonics	Fun	Emotion/Expression
26	The instrument offers new possibilities to express myself musically	Expressiveness	Hedonics	Fun	Emotion/Expression
27	I can express emotions through the instrument	Expressiveness	Hedonics	Fun	Emotion/Expression
28	I perceived the instrument as having its own character	Aliveness	Hedonics	Mysticism	
29	I can produce a warm atmosphere with the instrument	Aliveness	Hedonics	Mysticism	
30	I perceive the instrument as being alive	Aliveness	Hedonics	Mysticism	
31	I perceived the instrument as having some kind of soul	Aliveness	Hedonics	Mysticism	
32	Playing the instrument allows me to further develop my musical skills	Challenge	Motivation	Challenge	
33	I perceive the instrument as challenging in a positive way	Challenge	Motivation	Challenge	
34	Playing the instrument feels great	Comfort	Playing experience	Feel	
35	The instrument provides good feedback to me	Comfort	Playing experience	Feel	
36	The instrument suits my needs	Comfort	Playing experience	Feel	
37	The instrument allows me to focus on sound generation	Comfort	Playing experience	Playability	Control
38	Interacting with the instrument is surrounded by comfortable tactile feedback	Comfort	Playing experience	Feel	
39	The instrument suits my habits	Comfort	Playing experience	Usability	Habit
40	I perceive the instrument as playing along with me	Comfort	Playing experience	Feel	Personalization
41	I can control the sound appropriately	Control	Playing experience	Playability	Control
42	I feel in control of the instrument	Control	Playing experience	Playability	Control
43	I can play precisely on the instrument	Control	Playing experience	Playability	Control

Nr.	Item Text	Refined categories	Level 1 (a.d.)	Level 2 (a.d.)	Level 3 (a.d.)
44	I perceive the instrument as easy to control	Control	Playing experience	Playability	Control
45	I have to fight for getting what I want out of the instrument	Control	Playing experience	Playability	Control
46	The instrument processes my movements appropriately	Ergonomics	Playing experience	Playability	Body/Ergonomics
47	I feel comfortable when I play the instrument	Ergonomics	Playing experience	Feel	
48	I perceive the instrument as allowing small and efficient movements	Ergonomics	Playing experience	Playability	Body/Ergonomics
49	The instrument feels like an extension to my body	Ergonomics	Playing experience	Feel	Body Elements
50	The instrument appears ergonomically designed	Ergonomics	Playing experience	Playability	Body/Ergonomics
51	The instrument feels like a tailor-made suit	Ergonomics	Playing experience	Feel	Merging
52	The needed posture to play the instrument does not appeal to me	Ergonomics	Playing experience	Playability	Body/Ergonomics
53	The instrument does not bore me	Explorability	Motivation	Surprise	
54	I can continually discover new things by using the instrument	Explorability	Motivation	Freedom	
55	Producing sound with the instrument gives me a sense of achievement	Explorability	Motivation	Challenge	
56	The instrument offers me interesting possibilities to manipulate sound	Freedom	Motivation	Freedom	
57	The instrument offers me new possibilities of things to do	Freedom	Motivation	Freedom	
58	The instrument allows me to learn new things	Freedom	Motivation	Challenge	
59	The instrument offers me new facets of playing	Freedom	Motivation	Freedom	
60	I think the instrument offers the right amount of features to express sound	Freedom	Motivation	Freedom	
61	I can imagine applying the instrument in different musical contexts	Freedom	Motivation	Freedom	
62	The instruments expands my experience of musical interaction	Freedom	Motivation	Freedom	
63	I feel free when I play the instrument	Freedom	Motivation	Freedom	
64	The instrument feels like I can go beyond myself	Freedom	Motivation	Freedom	

Nr.	Item Text	Refined categories	Level 1 (a.d.)	Level 2 (a.d.)	Level 3 (a.d.)
65	I experience a lot of freedom using the instrument	Freedom	Motivation	Freedom	
66	Interacting with the instrument feels like crossing borders	Freedom	Motivation	Freedom	
67	I think the instrument is very well manufactured	build quality	Manufacturing	Tech. Manufacturing	
68	The optical cues facilitate my orientation on the instrument	Mapping	Playing experience	Usability	Examination/Understanding
69	It's easy for me to imagine possible actions because the good optical mapping	Mapping	Playing experience	Usability	Examination/Understanding
70	By looking at the instrument I can understand how it can be played	Mapping	Playing experience	Usability	Examination/Understanding
71	I think the instrument is innovative	Novelty	Motivation	Surprise	
72	I can recognize that the instrument responds well to my playing	Play Comfort	Playing experience	Feel	Self-induced causality
73	I perceive the instrument as responding dynamically	Play Comfort	Hedonics	Fun	Emotion/Expression
74	I perceive the instrument as comfortable to play	Play Comfort	Playing experience	Feel	
75	The material allows me to use comfortable gestures	Play Comfort	Playing experience	Playability	Body/Ergonomics
76	I think the material of the instrument supports the sound generation positively	Play Comfort	Manufacturing	Material	
77	The material feels pleasant to me	Play Comfort	Playing experience	Feel	
78	I feel relaxed when I play the instrument	Play Comfort	Playing experience	Feel	
79	The instrument feels clean to me	Play Comfort	Playing experience	Feel	
80	I think it is possible to do things with the instrument with little effort	Play Comfort	Playing experience	Playability	
81	I think the instrument is convenient to play with	Play Comfort	Playing experience	Feel	
82	Playing the instrument feels physically unpleasant	Play Comfort	Playing experience	Feel	Body Elements
83	The instrument pleases me sound-wise	Sound Quality	Sound	Quality	
84	I perceive the generated sounds as very good	Sound Quality	Sound	Quality	
85	I think the instrument offers a great dynamic range	Sound Quality	Sound	Quality	
86	The sound quality is convincing to me	Sound Quality	Sound	Quality	

Nr.	Item Text	Refined categories	Level 1 (a.d.)	Level 2 (a.d.)	Level 3 (a.d.)
87	The sound of the instrument is special to me	Sound Quality	Sound	Quality	
88	I think the instrument has an interesting timbre	Sound Quality	Sound	Timbre	
89	I perceive the sound of the instrument as warm	Sound Quality	Sound	Quality	
90	I can rely on the instrument when playing it	Stability	Manufacturing	Tech. Manufacturing	
91	I think the instrument is reliable	Stability	Manufacturing	Tech. Manufacturing	
92	I think the instrument is able to endure a lot	Stability	Manufacturing		
93	I perceive the instrument as solid	Stability	Manufacturing	Tech. Manufacturing	
94	I fear to break parts of the instrument when I play it	Stability	Manufacturing	Tech. Manufacturing	
95	I perceive the instrument as fragile	Stability	Manufacturing	Tech. Manufacturing	
96	Some parts of the instrument appear slack and baggy to me	Stability	Manufacturing	Tech. Manufacturing	
97	I perceive the instrument as sustainable in a musical way	Sustainability	Manufacturing	Tone/Quality	
98	I think I could maintain the instrument easily	Sustainability	Manufacturing		
99	I perceive the instrument as offering a lot of variety	Variety	Motivation	Freedom	
100	I have a good variety of sounds to choose	Variety	Motivation	Freedom	

Appendix B

Table B

Items ordered by descending mean importance (Study 2). Exclusion criteria; A: a relative majority of experts indicated a low importance for the item, B: (average rating < 4.0, thus falling below the neutral middle category, C: unclear or problematic item and phrasing issues due to e.g. negative phrasing, D: semantically redundant item. The columns 'Low', 'Mid' and 'High' contain the percentages of experts of the transformed 7-Point Likert scale to the three categories of agreement. Bold items were retained.

Nr	Item	Category	μ	(*SD*)	95% CI	PoA* 7	PoA** 3	Low %	Mid %	High %	High-Low Δ	Exclusion
31	**The instrument allows me to be creative**	Creativity	6.25	1.00	5.97 - 6.53	0.23	0.92	2.08	2.08	95.83	93.75	
22	**The instrument responds well to my actions**	Conformance	6.23	0.83	5.99 - 6.46	0.21	0.96	2.08	0.00	97.92	95.83	
90	**I can rely on the instrument when playing it**	Stability	6.21	0.90	5.95 - 6.46	0.20	0.88	0.00	6.25	93.75	93.75	
42	**I have fun playing the instrument**	Enjoyment	6.08	0.87	5.84 - 6.33	0.17	0.88	0.00	6.25	93.75	93.75	
53	**The instrument allows me to express myself**	Expressiveness	6.06	1.02	5.77 - 6.35	0.21	0.88	2.08	4.17	93.75	91.67	
25	**I can control the sound appropriately**	Control	6.04	0.99	5.76 - 6.32	0.16	0.84	2.08	6.25	91.67	89.58	
24	**I feel in control of the instrument**	Control	6.02	1.12	5.70 - 6.34	0.27	0.80	4.17	6.25	89.58	85.42	
83	**The instrument pleases me sound-wise**	Sound Quality	6.02	1.19	5.68 - 6.36	0.31	0.80	4.17	6.25	89.58	85.42	
96	**I think the instrument is reliable**	Stability	6.02	1.19	5.68 - 6.36	0.31	0.92	4.17	0.00	95.83	91.67	
39	**The instrument allows me to be engaged when I'm playing it**	Engagement	5.98	0.96	5.71 - 6.25	0.19	0.81	0.00	10.42	89.58	89.58	
32	**The instrument supports me in creating new music in any style**	Creativity	5.94	1.17	5.61 - 6.27	0.16	0.74	2.08	12.50	85.42	83.33	
30	**The instrument fosters my creativity**	Creativity	5.92	1.23	5.57 - 6.27	0.20	0.77	6.25	6.25	87.50	81.25	

Nr	Item	Category	μ	(*SD*)	95% CI	PoA[*] 7	PoA[**] 3	Low %	Mid %	High %	High-Low Δ	Exclusion
82	I perceive the generated sounds as very good***	Sound Quality	5.92	1.09	5.61 - 6.22	0.24	0.77	4.17	8.33	87.50	83.33	D
38	**I can imagine to play the instrument for a long time**	Engagement	5.85	0.87	5.61 - 6.10	0.19	0.88	0.00	6.25	93.75	93.75	
79	**I can recognize that the instrument responds well to my playing**	Play Comfort	5.85	1.13	5.53 - 6.17	0.16	0.92	4.17	0.00	95.83	91.67	
88	**I think the instrument offers a great dynamic range**	Sound Quality	5.81	1.23	5.46 - 6.16	0.15	0.65	4.17	16.67	79.17	75.00	
37	**I feel the urge to play the instrument again**	Engagement	5.79	1.32	5.42 - 6.17	0.20	0.81	8.33	2.08	89.58	81.25	
51	**The instrument does not bore me**	Explorability	5.77	0.90	5.51 - 6.03	0.28	0.88	2.08	4.17	93.75	91.67	
23	**I feel like I am initiating, executing and controlling the behavior of the instrument**	Conformance	5.71	1.20	5.37 - 6.05	0.19	0.74	2.08	12.50	85.42	83.33	
28	**I can play precisely on the instrument**	Control	5.67	1.26	5.31 - 6.02	0.33	0.80	4.17	6.25	89.58	85.42	
36	**The instrument keeps me interested**	Engagement	5.67	1.36	5.28 - 6.05	0.18	0.77	8.33	4.17	87.50	79.17	
40	**It is easy for me to get into a flow of playing with the instrument**	Engagement	5.65	1.28	5.28 - 6.01	0.20	0.70	6.25	10.42	83.33	77.08	
66	**The instrument offers me interesting possibilities to manipulate sound**	Freedom	5.63	1.30	5.26 - 5.99	0.17	0.70	8.33	8.33	83.33	75.00	
87	**The sound quality is convincing to me**	Sound Quality	5.63	1.62	5.17 - 6.08	0.25	0.67	10.42	8.33	81.25	70.83	
56	**The instrument offers me new possibilities of things to do**	Freedom	5.60	1.25	5.25 - 5.96	0.34	0.73	6.25	8.33	85.42	79.17	
84	**The sound of the instrument is special to me**	Sound Quality	5.60	1.54	5.17 - 6.04	0.19	0.61	10.42	12.50	77.08	66.67	
85	**I think the instrument has an interesting timbre**	Sound Quality	5.60	1.47	5.19 - 6.02	0.26	0.64	8.33	12.50	79.17	70.83	
12	**Playing the instrument feels great**	Comfort	5.58	1.33	5.21 - 5.96	0.25	0.73	6.25	8.33	85.42	79.17	

Nr	Item	Category	µ	(*SD*)	95% CI	PoA* 7	PoA** 3	Low %	Mid %	High %	High-Low Δ	Exclu sion
50	**I can continually discover new things by using the instrument**	Explorability	5.58	1.11	5.27 - 5.90	0.36	0.77	4.17	8.33	87.50	83.33	
17	**The instrument provides good feedback to me**	Comfort	5.56	1.22	5.22 - 5.91	0.19	0.80	4.17	6.25	89.58	85.42	
43	**The instrument processes my movements appropriately**	Ergonomics	5.48	1.20	5.14 - 5.82	0.22	0.68	4.17	14.58	81.25	77.08	
58	**The instrument allows me to learn new things**	Freedom	5.48	1.13	5.16 - 5.80	0.19	0.71	2.08	14.58	83.33	81.25	
41	**The instrument surprises me in a positive way**	Engagement	5.44	1.30	5.07 - 5.81	0.15	0.64	8.33	12.50	79.17	70.83	
55	**The instrument offers new possibilities to express myself musically**	Expressiveness	5.44	1.41	5.04 - 5.84	0.30	0.73	8.33	6.25	85.42	77.08	
80	**I perceive the instrument as responding dynamically**	Play Comfort	5.44	1.27	5.08 - 5.80	0.17	0.67	6.25	12.50	81.25	75.00	
13	**The instrument suits my needs**	Comfort	5.40	1.62	4.94 - 5.85	0.26	0.64	12.50	8.33	79.17	66.67	
11	**Playing the instrument allows me to further develop my musical skills**	Challenge	5.35	1.41	4.96 - 5.75	0.29	0.61	8.33	14.58	77.08	68.75	
49	**I feel comfortable when I play the instrument**	Ergonomics	5.33	1.26	4.98 - 5.69	0.28	0.67	6.25	12.50	81.25	75.00	
52	**Producing sound with the instrument gives me a sense of achievement**	Explorability	5.33	1.56	4.89 - 5.78	0.18	0.53	6.25	25.00	68.75	62.50	
54	**I can express emotions through the instrument**	Expressiveness	5.31	1.72	4.83 - 5.80	0.16	0.56	14.58	12.50	72.92	58.33	
59	**The instrument offers me new facets of playing**	Freedom	5.27	1.28	4.91 - 5.63	0.25	0.64	8.33	12.50	79.17	70.83	
44	**I perceive the instrument as allowing small and efficient movements**	Ergonomics	5.23	1.36	4.85 - 5.61	0.20	0.57	6.25	20.83	72.92	66.67	
75	**I perceive the instrument as comfortable to play**	Play Comfort	5.23	1.08	4.92 - 5.53	0.20	0.67	6.25	12.50	81.25	75.00	
29	**The instrument supports me in releasing my inner muse**	Creativity	5.21	1.89	4.67 - 5.74	0.17	0.65	16.67	4.17	79.17	62.50	
21	**The instrument does what I want it to do**	Conformance	5.19	1.67	4.71 - 5.66	0.22	0.62	16.67	6.25	77.08	60.42	

Nr	Item	Category	μ	(*SD*)	95% CI	PoA* 7	PoA** 3	Low %	Mid %	High %	High-Low Δ	Exclusion
60	I think the instrument offers the right amount of features to express sound	Freedom	5.19	1.20	4.85 - 5.53	0.30	0.57	6.25	20.83	72.92	66.67	C, D
19	**The instrument works the way I expect it to**	Conformance	5.17	1.65	4.70 - 5.63	0.28	0.47	14.58	20.83	64.58	50.00	
15	**The instrument allows me to focus on sound generation**	Comfort	5.10	1.53	4.67 - 5.54	0.25	0.46	10.42	27.08	62.50	52.08	
89	**I think the instrument is very well manufactured**	Stability	5.10	1.57	4.66 - 5.55	0.20	0.53	14.58	14.58	70.83	56.25	
98	**I perceive the instrument as sustainable in a musical way**	Sustainability	5.10	1.48	4.69 - 5.52	0.23	0.56	10.42	16.67	72.92	62.50	
100	**I perceive the instrument as offering a lot of variety**	Variety	5.08	1.46	4.67 - 5.50	0.23	0.54	12.50	16.67	70.83	58.33	
10	**I perceive the instrument as challenging in a positive way**	Challenge	5.06	1.44	4.66 - 5.47	0.22	0.59	10.42	14.58	75.00	64.58	
97	**I think I could maintain the instrument easily**	Sustainability	5.04	1.44	4.63 - 5.45	0.25	0.59	10.42	14.58	75.00	64.58	
63	**I can imagine applying the instrument in different musical contexts**	Freedom	4.98	1.60	4.53 - 5.43	0.22	0.49	14.58	18.75	66.67	52.08	
99	I have a good variety of sounds to choose	Variety	4.96	1.68	4.48 - 5.43	0.26	0.45	14.58	22.92	62.50	47.92	C, D
18	**Interacting with the instrument is surrounded by comfortable tactile feedback**	Comfort	4.94	1.54	4.50 - 5.37	0.16	0.54	12.50	16.67	70.83	58.33	
3	**I find the instrument aesthetically pleasing**	Aesthetics	4.92	1.30	4.55 - 5.28	0.26	0.59	12.50	12.50	75.00	62.50	
74	**The material allows me to use comfortable gestures**	Play Comfort	4.88	1.39	4.48 - 5.27	0.17	0.49	12.50	20.83	66.67	54.17	
94	**I think the instrument is able to endure a lot**	Stability	4.88	1.73	4.38 - 5.37	0.25	0.51	16.67	14.58	68.75	52.08	
65	**The instruments expands my experience of musical interaction**	Freedom	4.81	1.55	4.37 - 5.25	0.18	0.51	16.67	14.58	68.75	52.08	
81	**I think the material of the instrument supports the sound generation positively**	Play Comfort	4.81	1.45	4.40 - 5.22	0.27	0.47	12.50	22.92	64.58	52.08	
57	**I feel free when I play the instrument**	Freedom	4.79	1.64	4.33 - 5.25	0.30	0.43	12.50	29.17	58.33	45.83	
73	**The material feels pleasant to me**	Play Comfort	4.77	1.51	4.35 - 5.20	0.15	0.51	14.58	16.67	68.75	54.17	

Nr	Item	Category	μ	(*SD*)	95% CI	PoA* 7	PoA** 3	Low %	Mid %	High %	High-Low Δ	Exclusion
91	**I perceive the instrument as solid**	Stability	4.71	1.54	4.27 - 5.14	0.18	0.43	16.67	22.92	60.42	43.75	
9	**I perceived the instrument as having its own character**	Aliveness	4.67	1.43	4.26 - 5.07	0.28	0.42	16.67	25.00	58.33	41.67	
61	**The instrument feels like I can go beyond myself**	Freedom	4.67	1.79	4.16 - 5.17	0.30	0.47	22.92	12.50	64.58	41.67	
47	**The instrument feels like an extension to my body**	Ergonomics	4.63	1.94	4.08 - 5.17	0.26	0.45	29.17	10.42	60.42	31.25	
20	**I can use the instrument intuitively**	Conformance	4.60	1.63	4.14 - 5.07	0.16	0.49	27.08	8.33	64.58	37.50	
26	I perceive the instrument as easy to control	Control	4.54	1.49	4.12 - 4.96	0.20	0.40	20.83	22.92	56.25	35.42	D
78	**I feel relaxed when I play the instrument**	Play Comfort	4.46	1.60	4.01 - 4.91	0.21	0.37	22.92	25.00	52.08	29.17	
2	**I perceive the instrument as flawless**	Aesthetics	4.44	1.81	3.93 - 4.95	0.15	0.39	25.00	20.83	54.17	29.17	
64	I experience a lot of freedom using the instrument	Freedom	4.42	1.62	3.96 - 4.88	0.33	0.37	20.83	29.17	50.00	29.17	D
68	**The optical cues facilitate my orientation on the instrument**	Mapping	4.40	1.48	3.98 - 4.82	0.33	0.40	22.92	20.83	56.25	33.33	
46	**The instrument appears ergonomically designed**	Ergonomics	4.29	1.73	3.80 - 4.78	0.21	0.44	31.25	10.42	58.33	27.08	
1	**The instrument looks appealing**	Aesthetics	4.25	1.42	3.85 - 4.65	0.23	0.41	29.17	14.58	56.25	27.08	
4	**I perceive the instrument as unique**	Aesthetics	4.19	1.81	3.68 - 4.70	0.16	0.37	29.17	20.83	50.00	20.83	
67	It's easy for me to imagine possible actions because of the good optical mapping	Mapping	4.19	1.70	3.71 - 4.67	0.37	0.36	25.00	25.00	50.00	25.00	C, D
76	**The instrument feels clean to me**	Play Comfort	4.10	1.37	3.72 - 4.49	0.23	0.34	27.08	29.17	43.75	16.67	
77	**I think it is possible to do things with the instrument with little effort**	Play Comfort	4.10	1.59	3.66 - 4.55	0.24	0.33	31.25	27.08	41.67	10.42	
69	**By looking at the instrument I can understand how it can be played**	Mapping	4.08	1.71	3.60 - 4.57	0.32	0.36	35.42	18.75	45.83	10.42	

Nr	Item	Category	μ	(*SD*)	95% CI	PoA* 7	PoA** 3	Low %	Mid %	High %	High-Low Δ	Exclusion
71	I think the instrument is convenient to play with	Play Comfort	4.08	1.62	3.62 - 4.54	0.21	0.38	20.83	27.08	52.08	31.25	D
86	I perceive the sound of the instrument as warm	Sound Quality	4.06	1.81	3.55 - 4.58	0.26	0.33	27.08	33.33	39.58	12.50	C
33	**I perceive the instrument as easy to use**	Ease of Use	4.00	1.49	3.58 - 4.42	0.21	0.32	33.33	29.17	37.50	4.17	
7	I can produce a warm atmosphere with the instrument	Aliveness	3.94	2.16	3.33 - 4.55	0.37	0.40	41.67	10.42	47.92	6.25	B, C
35	**I can personalize the instrument**	**Ease of Use**	3.92	1.69	3.44 - 4.39	0.18	0.33	29.17	31.25	39.58	10.42	B
14	The instrument suits my habits	Comfort	3.85	1.79	3.35 - 4.36	0.18	0.32	35.42	29.17	35.42	0.00	B
45	The instrument feels like a tailor-made suit	Ergonomics	3.75	2.05	3.17 - 4.33	0.19	0.36	43.75	16.67	39.58	-4.17	A, B, C
70	I think the instrument is innovative	Novelty	3.75	1.87	3.22 - 4.28	0.24	0.39	41.67	12.50	45.83	4.17	B, C
16	I perceive the instrument as playing along with me	Comfort	3.73	1.87	3.20 - 4.26	0.26	0.35	39.58	18.75	41.67	2.08	B
34	I perceive the instrument as unnecessarily complex	Ease of Use	3.71	1.82	3.19 - 4.22	0.20	0.34	37.50	22.92	39.58	2.08	B, C
48	The needed posture to play the instrument does not appeal to me	Ergonomics	3.71	1.82	3.19 - 4.22	0.21	0.32	37.50	31.25	31.25	-6.25	A, B, C
62	Interacting with the instrument feels like crossing borders	Freedom	3.69	1.72	3.20 - 4.17	0.25	0.33	33.33	27.08	39.58	6.25	B
6	I perceive the instrument as being alive	Aliveness	3.67	2.12	3.07 - 4.27	0.18	0.38	50.00	14.58	35.42	-14.58	A, B, C
93	I fear to break parts of the instrument when I play it	Stability	3.65	2.22	3.02 - 4.27	0.17	0.42	47.92	8.33	43.75	-4.17	A, B, C
72	Playing the instrument feels physically unpleasant	Play Comfort	3.63	2.23	2.99 - 4.26	0.14	0.37	47.92	16.67	35.42	-12.50	A, B, C
92	I perceive the instrument as fragile	Stability	3.48	1.87	2.95 - 4.01	0.15	0.37	45.83	16.67	37.50	-8.33	

Nr	Item	Category	μ	(*SD*)	95% CI	PoA[*] 7	PoA[**] 3	Low %	Mid %	High %	High-Low Δ	Exclu sion
8	I perceived the instrument as having some kind of soul	Aliveness	3.40	2.02	2.82 - 3.97	0.26	0.36	47.92	20.83	31.25	-16.67	A, B
27	I have to fight for getting what I want out of the instrument	Control	3.40	1.81	2.88 - 3.91	0.22	0.36	47.92	20.83	31.25	-16.67	A, B
95	Some parts of the instrument appear slack and baggy to me	Stability	3.38	1.81	2.86 - 3.89	0.26	0.34	39.58	37.50	22.92	-16.67	A, B, C
5	The colors of the instrument are appealing	Aesthetics	2.96	1.41	2.56 - 3.36	0.20	0.45	62.50	16.67	20.83	-41.67	A, B

* *Proportion of agreement for the original 7-Point Likert-Scale (ranging from 0 to 1, with 1 = 100% agreement, not corrected for agreement by chance).*

** *Proportion of for the compressed three categories (ranging from 0 to 1, with 1 = 100% agreement, not corrected for agreement by chance).*

*** *Item Nr. 82 was replaced with a better phrased item (Nr. 84, Appendix A).*