Experience of Groove Questionnaire: Instrument Development and Initial Validation

Olivier Senn, Toni Bechtold, Dawn Rose
Lucerne University of Applied Sciences and Arts, Lucerne, Switzerland

Guilherme Schmidt Câmara
University of Oslo, Oslo, Norway

Nina Düvel
Hanover University of Music, Drama and Media, Hanover, Germany

Rafael Jerjen, Lorenz Kilchenmann, Florian Hoesl, Antonio Baldassarre, & Elena Alessandri
Lucerne University of Applied Sciences and Arts, Lucerne, Switzerland

Music often triggers a pleasurable urge in listeners to move their bodies in response to the rhythm. In music psychology, this experience is commonly referred to as *groove*. This study presents the *Experience of Groove Questionnaire*, a newly developed self-report questionnaire that enables respondents to subjectively assess how strongly they feel an *urge to move* and *pleasure* while listening to music. The development of the questionnaire was carried out in several stages: candidate questionnaire items were generated on the basis of the groove literature, and their suitability was judged by fifteen groove and rhythm research experts. Two listening experiments were carried out in order to reduce the number of items, to validate the instrument, and to estimate its reliability. The final questionnaire consists of two scales with three items each that reliably measure respondents' *urge to move* (Cronbach's $\alpha = .92$) and their experience of *pleasure* ($\alpha = .97$) while listening to music. The two scales are highly correlated ($r = .80$), which indicates a strong association between motor and emotional responses to music. The scales of the *Experience of Groove Questionnaire* can independently be applied in groove research and in a variety of other research contexts in which listeners' subjective experience of music-induced movement and enjoyment need to be addressed: for example the study of the interaction between music and motivation in sports and

research on therapeutic applications of music in people with neurological movement disorders.

THE TERM *MUSICAL GROOVE* WAS ORIGINALLY coined within the African American popular music communities (Danielsen, 2006; Doffman, 2008; Roholt, 2014). It is frequently applied to funk, soul, rock, R&B, jazz, rap, and other musical styles that, according to Pressing (2002, p. 285), derive from the "African diasporic" performance tradition. These genres share many rhythmic and stylistic properties (Câmara & Danielsen, 2018). In this context, *groove* has a variety of meanings. It denotes aspects of the structure, performance, and perception of music: the noun *a groove* may refer to a repetitive musical pattern that functions as the basis of a track of popular music, particularly in funk, soul, rock, or post-1960 jazz (Kernfeld, 2003; Zbikowski, 2004). The verb *to groove* means playing music together in an effortless and rhythmically well-coordinated manner (Berliner, 1994, p. 349; Butterfield, 2011; Keil, 2010; Monson, 1996, p. 67). Finally, when popular music *has groove*, it is aesthetically pleasing and invites listeners to participate with body movements such as clapping, head-bobbing, swaying, stomping, or dancing (Iyer, 2002; Roholt, 2014).

Madison (2006) was the first study to develop groove as a psychological concept. Music psychology defines the *experience of groove* more narrowly as listeners' inner urge to move in response to the music (Davies, Madison, Silva, & Guyon, 2013; Eaves, Griffiths, Burridge, McBain, & Butcher 2019; Madison, 2001, 2006; Madison & Sioros, 2014; Sioros, Miron, Davies, Gouyon, & Madison, 2014). Many studies additionally describe this inner urge as being pleasant (Cameron et al., 2019; Etani, Marui, Kawase, & Keller, 2018; Janata, Tomic, & Haberman, 2012; Lustig & Tan, 2019; Madison, 2003; Madison, Gouyon, Ullèn, & Hörnström, 2011; Matsushita & Nomura, 2016; Matthews, Witek, Heggli, Penhune, & Vuust, 2019; Senn, Kilchenmann, Bechtold,

& Hoesl, 2018; Senn, Kilchenmann, von Georgi, & Bullerjahn, 2016; Stupacher, Hove, & Janata, 2016; Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014).

Some psychological studies present more complex definitions of groove (Frühauf, Kopiez, & Platz, 2013; Hosken, 2020; Wesolowski & Hofmann, 2016; Witek, 2016). These composite definitions of the groove experience show similarities to the multifaceted definitions used by musicians and music listeners, but they also emphasize the relevance of body movement and pleasure. To summarize, the definition of groove as music listeners' inner urge to move their bodies in response to the music is unanimously accepted in music psychology, but a great majority of studies also point out that groove is associated with an experience of pleasure.

The psychological groove experience is not defined by specific musical or stylistic features. Consequently, its scope can be expanded beyond Western popular music. It is transcultural in the sense that the experience may be triggered by music from any style and cultural background: whenever people feel a pleasurable urge to move with any kind of music, they experience groove as defined by music psychology.

Psychological groove research investigates the causes of the groove experience. In the past, researchers have studied many factors that potentially influence the intensity of the experience: Research has focused on musical properties such as rhythmic and harmonic complexity, beat salience, microtiming, tempo, or musical patterns. Factors linked to the person of the listener have also been investigated; for example, listeners' taste, musical expertise, and background. Finally, the concrete listening situation (such as live vs. recorded music) and the interaction of auditory, visual, and tactile modalities have also been studied (for a recent overview, see Eaves et al., 2019; Hove, Martinez, & Stupacher, 2020; Matthews et al. 2019; Senn, Rose, et al., 2019, Swarbrick et al., 2019).

In order to study the causes of the groove experience, it is of paramount importance to be able to assess and quantify the intensity with which groove is experienced by listeners: groove needs to be measurable. Previous research has used motion capture technology to measure the amplitude and synchronicity of listeners' body movement as a direct physical expression of groove (see Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2013; Danielsen, Haugen, & Jensenius, 2015; Kilchenmann & Senn, 2015; Swarbrick et al., 2019). Neuroscientific approaches have measured the entrained neural oscillations in the brain of a listener as an indicator that the temporal regularities of the music have been decoded in perception (see Cameron et al., 2019; Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan,

Schönwiesner, Keller, Lenc, & Lehmann, 2018; Stupacher, Witte, Hove, & Wood, 2016).

In the majority of behavioral groove studies, however, listeners rated the intensity of their groove experience via a self-report questionnaire. A variety of single-item Likert scales has been used to measure groove as an urge to move or as pleasure (see these operationalizations in Tables A and B in the Appendix).

The use of many different psychometric scales for measuring listeners' subjective groove experience is problematic for two reasons: First, scales that use different wordings cannot be assumed to measure exactly the same underlying construct, even if the scales have the same name (such as *urge to move* or *pleasure*). Consequently, the comparison of effects across studies is difficult. Second, most of the scales have not been subjected to a systematic procedure of psychometric validation. It is therefore unclear to what extent these scales reliably measure the underlying psychological construct. (An exception is the *Emotional Assessment of Groove* questionnaire by von Georgi, Bullerjahn, Senn, & Kilchenmann, 2016, which has been validated, but is not congruent to the two-dimensional definition of groove as pleasurable urge to move in response to music.)

The goal of this study is to provide the groove research community with the *Experience of Groove Questionnaire*, a newly developed psychometric instrument. The questionnaire operationalizes the two theoretical definitions of groove that are frequently used today. It consists of two valid, reliable, and concise multi-item scales. One is dedicated to measuring the subjective impression of an *urge to move* in the listener, and the other assesses the subjective experience of *pleasure* while listening. Since opinions diverge whether *pleasure* is essential to groove or not, the two scales need to be applicable independently of one another.

The availability of a standard questionnaire with known psychometric properties will simplify the design of future groove studies. The consistent application of validated measures makes it possible to compare results across groove studies. Multi-item scales are not only more reliable than single-item scales (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012), they also open a path to complex models of groove that involve mediation, moderation, and indirect effects (see the model proposed by Senn, Rose, et al., 2019).

Method and Materials

The development of the *Experience of Groove Questionnaire* observed the design principles and best practice recommendations proposed in McCoach, Gable, and

Madura (2013). Their book, entitled *Instrument Development in the Affective Domain*, proposes a comprehensive methodology for questionnaire design and validation. It starts from the theoretical definition of the construct of interest, moves on to the operationalization of questionnaire items, to the selection of questionnaire items and finally to the validation of the resulting scales. McCoach et al. (2013) describe factor analysis techniques only sparingly. Consequently, several complementary treatises on factor analysis were consulted for this study (Bollen, 1989; Brown, 2006; Gorsuch, 2015; Pett, Lackey, & Sullivan, 2003; Thompson, 2004).

The general procedure was as follows:

→ A set of 25 candidate questionnaire items to measure the two constructs (*urge to move* and *pleasure*) were derived from the operational definitions given in the groove literature (see Tables A and B in the *Appendix*). Some items were additionally formulated by the authors. The wording of the items was limited to Ogden's Basic English vocabulary (Ogden, 1948; Templer, 2007). This made sure that respondents with basic English skills and without music training could fill out the questionnaire.

→ The suitability of the candidate items was judged by 15 groove and rhythm research experts (Pre-Test). They rated: a) how well each item captured one of the two constructs, b) how relevant it was to the description of this construct; and c) how certain they were about their judgement. The five items deemed the least appropriate were subsequently discarded. Two further items that had been proposed by the experts were added to the list.

→ In a first listening experiment (Experiment 1) 56 listeners rated eight musical stimuli from different styles (extracted from commercially available recordings) using the 22 remaining candidate questionnaire items. The eight stimuli were chosen such that we expected them to elicit a range of reactions with respect to the *urge to move* and *pleasure*. Exploratory Factor Analysis revealed the factor structure of the questionnaire items, and indicated which items best captured the two underlying constructs.

→ Three items were selected for each construct that strongly correlated with the underlying latent construct and addressed clearly differentiated topics related to the construct.

→ In a second listening experiment (Experiment 2) a separate, independently recruited sample of 197 participants rated the same eight stimuli using the three selected items per scale. The reliability of the *urge to move* and *pleasure* scales was established with Confirmatory Factor Analysis, and their construct validity was assessed.

Details of the three-part process to develop the *Experience of Groove Questionnaire* are described in the following sections.

Pre-test: Content Validity of 25 Questionnaire Items

Previous studies have operationalized the urge to move and the experience of pleasure in a variety of ways. These operationalizations are listed in the Appendix as Tables A and B. The tables show that definitions present themselves as variations on a common theme, although there is considerable diversity in the specific wording.

By drawing upon this wealth of operational definitions, and adding several of our own, we generated 25 candidate questionnaire items (Table 1). Our goal was that this selection of items represents the "domain of content" (McCoach et al., 2013, p. 68) of the two underlying *urge to move* and *pleasure* constructs in a comprehensive way. This means that the questionnaire items approached these constructs from a variety of perspectives. We then asked groove and rhythm research experts to rate the content validity of the items in an online survey. Items were considered to have good content validity if they were relevant and appropriate to measure the underlying constructs of interest (Bollen, 1989, p. 185).

METHOD

Participants. Fifteen individuals (six female) participated in this survey. They were chosen because they were researchers with experience in the study of groove, entrainment, or rhythm perception. Each participant was active in academia (six professors, seven postdoc/senior researchers, two PhD students). They were affiliated with academic institutions in Germany ($n = 4$), the United Kingdom ($n = 3$), Japan ($n = 2$), Austria, Finland, Norway, Portugal, Sweden, and the United States ($n = 1$ each).

Procedure. Participants were invited by e-mail to fill in the online questionnaire, which was implemented on the *SosciSurvey* platform (www.soscisurvey.de). They received an access code to enter the survey and gave informed consent. Participants were presented a list of 25 candidate items (Table 1) in a random order. They assessed the content validity of each item by answering the questions specified in Table 2.

TABLE 1. *Candidate Questionnaire Items with Mean (SD) Content Validity Ratings for the Urge to Move and Pleasure Constructs*

		Content Validity	
Item No.	Item Text	Urge to Move	Pleasure
1	This music evokes the sensation of wanting to move some part of my body.	**.77** (.32)	.03 (.13)
2	I can feel the rhythm of this music in my body.	**.74** (.33)	.03 (.08)
3	This music makes me want to tap along.	**.61** (.32)	.06 (.16)
4	This music is good for dancing.	**.60** (.35)	.10 (.17)
5	This music makes me want to move.	**.57** (.38)	.13 (.18)
6	I cannot sit still while listening to this music.	**.55** (.30)	.03 (.10)
7	I feel like bobbing my head along with this music.	**.55** (.29)	.07 (.18)
8	I would like to dance to this music.	**.44** (.28)	.19 (.20)
9	I would like to listen to this music to accompany repetitive work.	**.14** (.21)	.01 (.03)
10	Listening to this music gives me pleasure.	.01 (.03)	**.71** (.35)
11	This music makes me feel good.	.02 (.04)	**.59** (.40)
12	Listening to this music is enjoyable.	.00 (.00)	**.47** (.38)
13	I love listening to this music.	.01 (.03)	**.37** (.36)
14	This music gives me positive feelings.	.03 (.06)	**.36** (.27)
15	I like listening to this music.	.03 (.10)	**.32** (.31)
16	This music makes me smile.	.00 (.00)	**.30** (.38)
17	My mood gets better when I am listening to this music.	.00 (.00)	**.26** (.19)
18	This music is fun to listen to.	.03 (.07)	**.24** (.26)
19	This music gives me positive energy.	.09 (.13)	**.18** (.13)
20	I feel energized by this music.	.16 (.14)	.16 (.12)
21	This is great music for a dance party.	**.38** (.29)	.13 (.18)
22	This is good music for gym/workout.	**.22** (.15)	.05 (.08)
23	This music evokes happy feelings.	.01 (.05)	**.43** (.35)
24	I want to hear more of this music.	.03 (.06)	**.11** (.15)
25	This music is interesting.	.00 (.00)	**.02** (.06)

Note: Items No. 21-25 were subsequently excluded due to low content validity ratings or substantial criticism by the respondents.

TABLE 2. *Quantitative Encoding of Respondents' Content Validity Ratings Using Numerical Weights*

		Weight	
Question	Possible answer	Urge to Move	Pleasure
Which construct is captured by the item?	Urge to Move	1.0	0.0
	Pleasure	0.0	1.0
	Both	0.5	0.5
	Neither	0.0	0.0
How relevant is the item for capturing this construct?	Very relevant	1.0	1.0
	Relevant	0.5	0.5
	Not relevant	0.0	0.0
How certain are you that the item captures this construct?	Certain	1.0	1.0
	Slightly certain	0.5	0.5
	Uncertain	0.0	0.0

(1) Participants were asked whether the item captured the *urge to move* or the *pleasure* concept, or both, or neither.
(2) They indicated how certain they were regarding their categorisation.
(3) They judged how relevant the item was for the chosen construct(s).

Participants could also comment on the items' wording and appropriateness in a free text field. They had the opportunity to propose further items if they thought that important topics were not represented in the list of items. Participants rated the candidate questionnaire items without listening to any musical stimuli.

Quantitative encoding of participant feedback. For each item, a composite numeric content validity rating in the range between zero and one was calculated from participants' feedback. This was carried out by multiplying the corresponding numerical weights of Table 2. For example, a participant might judge that an item captured the concept of *pleasure* (1.0) but not the *urge to move* (0.0), that the item was *relevant* (0.5) but not *very relevant* for measuring *pleasure*, and the participant was *certain* (1.0) that her assignment was correct. In this case, *pleasure* obtained a composite content validity rating of $1.0 \times 0.5 \times 1.0 = 0.5$, while the *urge to move* received a rating of $0.0 \times 0.5 \times 1.0 = 0$.

RESULTS AND DISCUSSION

Table 1 shows mean content validity ratings for the 25 items and the two constructs, *urge to move* and *pleasure* (standard deviations in parentheses). The list is presented in two blocks: participants assigned items 1–9 quite consistently to the *urge to move* construct, and items 10–20 to the *pleasure* construct. Within the blocks in Table 1, the items are ordered according to the content validity rating with respect to the dominant construct (bold script). Items with clear association and relevance to a construct appear on top of each block.

The pre-test indicated that items 21–25 were inappropriate for the questionnaire, and we listed them in a third block in Table 1. One respondent pointed out in the survey's free text section that items 21 and 22 referred to activities (party, gym) that are frequently encountered in Western societies, but not necessarily in other cultural contexts. Therefore, these items might show cultural bias. Another respondent objected to using the concept of happiness in the questionnaire (item 23): happiness refers to an idea of general well-being and satisfaction. The respondent argued that this is too grand a concept to capture a short-term impression about music, and we agreed with this argument. Also, the adjective "happy" is often used to describe the emotional content of the music itself (see Zentner, Grandjean, & Scherer, 2008), which might be a source of confusion. Finally, items 24–25 had very low content validity ratings for either of the constructs.

CONCLUSIONS

Items 1–20 were selected as candidate items for the first listening experiment (Experiment 1, below). Items 21–25 were dropped from the investigation for the reasons outlined above. Participants proposed two further candidate items ("This music puts a spring in my step" and "This music is good for physical activities") that were subsequently included as candidate items in Experiment 1.

Experiment 1: Exploration of 22 Questionnaire Items

A listening experiment was carried out using the 22 candidate questionnaire items (items 1–20 from Table 1 plus the two additional items). In this experiment, participants listened to eight musical stimuli and subsequently expressed their agreement with the items using seven-point Likert response scales. The goal of the experiment was to investigate the factor structure of the items.

METHOD

Stimuli. Stimuli consisted of eight musical excerpts of two minutes duration from commercially available recordings. They were selected from a wide variety of styles (Western and non-Western) with the purpose to trigger extreme reactions in both the *urge to move* and *pleasure* dimensions. In a first step, the second author compiled a list of 41 musical excerpts from his personal music collection. Eight stimuli were selected that we anticipated would elicit either low or high *urge to move* and either low or high *pleasure*. We selected two stimuli for each of the four feature combinations (see Table 3). A complete discography is presented in the Appendix. The classification of stimuli into high/low *urge to move* or *pleasure* groups represents the authors' expectation of how a general population participating in the listening experiment is likely to react to the music.

We expected two of the stimuli to score highly on both the *urge to move* and *pleasure*:

(1) "Superstition" is an R&B and Soul classic from Stevie Wonder's 1972 album "Talking Book." The track obtained high groove ratings in the seminal study by Janata et al. (2012). Accordingly, we expected many listeners to judge the song as triggering both a strong urge to move and pleasure.

(2) The track "Bala" from the 2007 album "Segu Blue" is centred on the Ngoni, a West African lute played by Bassekou Kouyate, who was accompanied by his band Ngoni Ba from Mali. We judged the rhythm of this piece to trigger an urge to move; and we also considered the music to be pleasing.

We predicted two stimuli would generate a strong *urge to move*, but low *pleasure*:

(3) Cash & Carry's "Tchip Tchip" (also known as "Chicken Dance") from the 1973 album with the same title is played entirely on synthesizers. It is a piece that (at least in German speaking

TABLE 3. *Stimuli for Listening Experiment 1*

No.	Title	Act/Composer	Urge to Move	Pleasure
1	Superstition	Stevie Wonder	high	high
2	Bala	Bassekou Kouyate	high	high
3	Tchip Tchip	Cash & Carry	high	low
4	Hamdouchi	M. Guinia / P. Sanders	high	low
5	Sunrise	Angelite & Moscow Art Trio	low	high
6	Symphony No. 3 (4th mov.)	Gustav Mahler	low	high
7	Machine Gun	Peter Brötzmann Octet	low	low
8	My Pal Foot Foot	The Shaggs	low	low

Note: The categorical *Urge to Move* and *Pleasure* variables specify our expectation regarding the ratings these stimuli were likely to obtain in a listening experiment.

countries) has a connotation of low-brow musical comedy. It is associated with a simple dance choreography. We judged that listeners would positively associate this music with body movement, but they would experience little pleasure because of the low prestige of the music and the squawky sound of the synthesizer.

(4) On "Hamdouchi," gnawa (lute) player Maleem Mahmoud Guinia from Morocco collaborated with American free jazz saxophonist Pharoah Sanders for the 1994 album "The Trance of Seven Colors." The music has a percussion track with an infectious rhythm. The wind instruments (saxophone and rhaita, a North African double reed instrument similar to a shawm) play their melodies with an intense, nasal timbre and many dissonant multiphonics. Our impression was that the music invites movement, but the sound textures may be unpleasant to listeners unfamiliar with the music.

For two stimuli, we assumed that they would trigger little *urge to move*, but high *pleasure*:

(5) Angelite is a Bulgarian women's choir. Their track "Sunrise" from the 1998 album "Mountain Tale" was recorded in collaboration with the Moscow Art Trio and Mongolian singer Huun-Huur-Tu. The meter is frequently suspended, and the regular pulse is interrupted, which makes it difficult to map synchronized body movement to the music. So, we assumed the music would trigger little impulse to move in listeners. Yet, we expected that respondents would have pleasure listening to the lush vocal harmony.

(6) The fourth movement of Gustav Mahler's Symphony No. 3 is a slow symphonic movement. In the excerpt, the string section of the orchestra presents an expressive melody. We used the 1989 recording of the New York Philharmonic

Orchestra, conducted by Leonard Bernstein. We assumed that the slow tempo and expansive tempo rubato would not trigger an urge to move in listeners, but that many would enjoy the rich timbre of the strings.

Finally, we chose two stimuli that we estimated would generate both little *urge to move* and little *pleasure* in listeners:

(7) The extract from "Machine Gun" by the Peter Brötzmann Octet's 1968 eponymous album is a high-energy free jazz performance, in which every band member appears to play as loud as possible. No common and regular time organisation (such as a meter) is detectable. We hypothesized that "Machine Gun" would neither motivate listeners to dance nor trigger much pleasure in the average listener.

(8) The Shaggs were a rock band active between 1968 and 1975. The band consisted of three (later four) sisters, whose father was convinced they would rise to stardom. He insisted on them forming a band despite the fact that they had neither the skill nor inclination (Chusid, 2000, p.3). The Shaggs' music is characterized by out-of-tune singing, discordant guitar playing, and very loose rhythmic interaction. "My Pal Foot Foot" is a song about the Shaggs' family cat from their album "Philosophy of the World" (1969). O'Connor (1996) hailed the album as "the record that renders all future incompetency irrelevant" (p. 61). We assumed that "My Pal Foot Foot" would create little urge to move and little pleasure in the average listener.

A two-minute excerpt was chosen from each piece in order to give participants enough time to rate the 22 candidate items while the music was playing. As far as possible, the two-minute excerpts were chosen to be

TABLE 4. *Pattern Matrix of the 22 Questionnaire Items (2 Correlated Factors, Promax Rotation)*

No.	Item	Urge to Move	Pleasure
1	This music is good for dancing.	**0.98**	−0.11
2	I cannot sit still while listening to this music.	**0.92**	−0.13
3	This music makes me want to move.	0.90	−0.03
4	This music evokes the sensation of wanting to move some part of my body.	**0.90**	−0.01
5	I feel like bobbing my head along with the music.	0.89	0.01
6	I would like to dance to this music.	0.86	0.04
7	This music makes me want to tap along.	0.84	0.05
8	This music is good for physical activities.	0.81	0.04
9	I can feel the rhythm of this music in my body.	0.77	0.07
10	This music puts a spring in my step.	0.67	0.23
11	I like listening to this music.	−0.14	**1.07**
12	I love listening to this music.	−0.09	1.00
13	Listening to this music gives me pleasure.	−0.05	**0.99**
14	Listening to this music is enjoyable.	−0.01	0.96
15	This music makes me feel good.	0.08	**0.89**
16	My mood gets better when I am listening to this music.	0.17	0.80
17	This music evokes positive feelings.	0.17	0.79
18	This music gives me positive energy.	0.22	0.73
19	This music is fun to listen to.	0.23	0.70
20	I feel energized by this music.	0.28	0.65
21	This music makes me smile.	0.25	0.62
22	I would like to listen to this music to accompany repetitive activities.	0.40	0.43

Note: Factor coefficients in bold print indicate items that were subsequently chosen to be used in Experiment 2.

relatively uniform in terms of tempo, instrumentation, dynamics, and mood. The original audio files were obtained from CD or purchased on iTunes. The loudness was adjusted to equal levels (average-weighted RMS at -22 dB) for each excerpt using *Audacity* (version 2.2.2); short fade-ins (1 s) and fade-outs (3 s) were added, and the music was exported in stereo to mp3 format (192 kbit/s).

Participants. A total of 56 participants (18 female, mean age 41 years, ranging from 22 to 68) were recruited. Invitations to participate were issued via e-mail to people who had participated in an earlier groove study and had given permission to be contacted again when a new survey was available. Several groove research experts who had participated in the Pre-Study helped distribute the invitation, but they did not participate in the experiment themselves. Participants self-identified as either professional musicians ($n = 17$), amateur musicians ($n = 20$), or music listeners ($n = 19$). Participants lived in Germany ($n = 23$), Switzerland ($n = 21$), Japan ($n = 7$), Canada ($n = 2$), Austria, Netherlands, and USA ($n = 1$ each).

Procedure. The listening experiment was implemented online via the *SosciSurvey* platform. Participants completed the survey once they had given informed consent. They answered a series of questions related to

demographic data (age, musical expertise, gender, country of residence, musical taste; for details, see Senn et al., 2018). They were instructed to use quality headphones or external loudspeakers during the experiment and to carry the experiment out in a quiet room. Participants listened to a test stimulus in order to adapt the playback loudness to a fairly loud, but comfortable level, and they were asked not to change the loudness level for the remainder of the experiment.

The eight experimental stimuli were presented in a randomized order. Participants assessed the degree to which they agreed with the 22 candidate items (Table 4) while listening to a stimulus, using a 7-point Likert scale (*strongly disagree* = 0, *disagree* = 1, *slightly disagree* = 2, *neither agree nor disagree* = 3, *slightly agree* = 4, *agree* = 5, *strongly agree* = 6). For each stimulus, candidate items were listed in a fixed order on the same screen. The fixed list had previously been randomized once, so items pertaining to the urge to move and pleasure dimensions alternated. Participants could start giving their ratings as the music started playing, and they could update their ratings if later events in the music changed their opinion. When participants were satisfied with their ratings, they could press the "Next" button in order to proceed to the next screen. If ratings were missing, participants were prompted to provide the missing ratings. When all ratings were

complete, pressing the "Next" button transmitted the ratings to the *SosciSurvey* database, and the participants proceeded to the next screen (next stimulus, or the end of the survey). On average, the experiment took 20 minutes to complete.

The choice of a 7-point Likert scale follows the recommendations of McCoach et al. (2013). They argue that respondents are familiar with Likert scales (p. 38) and that these scales produce consistent data (p. 48). Following arguments by Dolan (1994), DiStefano (2002) and Beauducel and Herzberg (2006), McCoach et al. (2013, p. 69) advocate for five to seven answer categories because these numbers of choices do not overwhelm the respondents, and because they produce sufficiently differentiated data that can be used for parametric statistical analysis (particularly, when several items are combined to form one scale).

Analysis. With 56 participants and eight stimuli, we collected $56 \times 8 = 448$ ratings on each of the 22 questionnaire items. We studied the associations between the items using exploratory factor analysis. The total number of observations was sufficiently large to carry out the analysis (Gorsuch, 2015, p. 350). The 22 questionnaire items had specifically been selected to measure listeners' *urge to move* or *pleasure*. Accordingly, we defined the target model to have two factors. The number of factors was confirmed by parallel analysis (McCoach et al., 2013, p. 123), Kaiser's criterion (Kaiser, 1958), and a scree plot. Based on the results of earlier studies (Matthews et al., 2019), we expected the two factors to be positively correlated. Accordingly, an oblique *promax* rotation was used. The analysis was carried out in *R* (version 3.5.1) and *R-Studio* (version 1.1.463), using the *factanal* function.

RESULTS

The pattern matrix of the exploratory factor analysis is presented in Table 4. It is ordered in two blocks depending on whether the pattern coefficient of the *urge to move* (top block) or *pleasure* (bottom block) factor is greater for any given item. Within each block, items are ordered with decreasing pattern coefficient of the dominant factor.

The two factors accounted for 73% of the variance across all items. This indicates that the items have high communality and show strong correlations with the two underlying concepts. Most items had only one single pattern coefficient with a high absolute value, so most items can unambiguously be assigned to one of the two concepts (the exception being item 22). As predicted, the *urge to move* and *pleasure* factors were strongly

associated in the context of Experiment 1. The interfactor correlation coefficient was high at $r = .75$.

The next step was to select items for the final questionnaire. We aimed to choose a small number of items (a minimum of three items per construct) that reliably measure the underlying constructs of *urge to move* and *pleasure*. Items that were listed in the top half of either block in Table 4 were most likely to be reliable indicators for the respective constructs. Further, items should also have high content validity (see experts' ratings in Table 2) and cover a wide content domain. This means that each item should focus on a different aspect of the respective construct: together, items should address the most important contexts within which the construct applies.

For the *urge to move* scale, we selected items 1, 2, and 4:

→ *Item 1* ("This music is good for dancing") addresses the danceability of the music. Dancing is a typical bodily expression of the urge to move. We preferred item 1 to the very similar item 6 ("I would like to dance to this music"), because it is more highly correlated with the *urge to move* factor, and it does not assume that the respondent personally likes dancing or is capable of dancing. A respondent can judge whether music is "good for dancing" even if she/he would not "like to dance" to it. Item 1 also better reflects the situation of the listening experiment, where respondents are normally seated at a computer and unable to get up and dance. Item 1 had a high mean content validity score of .60 (Table 1) and a high pattern coefficient of 0.98 (Table 4) with respect to the *urge to move* construct.

→ *Item 2* ("I cannot sit still while listening to this music") also had an over-average mean content validity score of .55 and a high pattern coefficient of 0.92 on *urge to move*. Item 2 addresses the construct from a completely different perspective than item 1 by accentuating an inner compulsion ("cannot sit still") that specifically captures the *urge* aspect of the construct.

→ *Item 4* ("This music evokes the sensation of wanting to move some part of my body") had a high mean content validity score of .77, and a high pattern coefficient of 0.90 on the *urge to move* factor. We chose item 4 over the similar item 3 because it obtained a higher mean content validity rating by the groove research experts, and its wording is close to earlier operational definitions of groove (see Davies et al., 2013; Madison, 2006; Madison et al., 2011; Madison & Sioros, 2014).

For the *pleasure* scale, we chose items 11, 13, and 15:

→ *Item 11* ("I like listening to this music") offers a hedonic assessment of the listening experience. The item was rated relatively low on content validity (.32) by the groove research experts but had a very high pattern coefficient of 1.07 on *pleasure*. Potentially, the groove research experts considered the statement to be too close to an aesthetic judgment, instead of an indicator of a hedonic experience. Item 12 ("I love listening to this music.") was not selected, because it was very similar to item 11 and we considered it to make too strong a claim that might result in a floor effect.

→ *Item 13* ("Listening to this music gives me pleasure") prompts listeners to directly assess their experience of pleasure. This item was rated high on content validity (.71) by the experts of the Pre-Test. It had a high pattern coefficient of 0.99 on *pleasure* and was slightly better aligned with the underlying factor than the very similar item 14 ("Listening to this music is enjoyable."). Item 13 is compatible with items used by Witek et al. (2014) and Matthews et al. (2019).

→ *Item 15* ("This music makes me feel good") focuses on participants' perception of their own mood and feeling. It obtained high mean content validity ratings (.59) and a high pattern coefficient (0.89) on *pleasure*.

All six chosen items showed high pattern coefficients with regard to one of the dominant constructs. Most of them obtained high mean content validity ratings and represent different perspectives on either of the two constructs, the *urge to move* and *pleasure*. Thus, they covered much of the content domain, and they were strongly correlated with the two factors emerging from the 22 items. We concluded that the resulting scales were likely to have high content validity.

Confirmatory factor analysis was carried out in *R* using the *Lavaan* (v. 0.6-3) package in order to assess the reliability of the six items as indicators of the two underlying constructs. The target model consisted of two correlated latent variables: *urge to move* (measured by items 1, 2, and 4) and *pleasure* (measured by items 11, 13, and 15). For comparison, nested models with two uncorrelated latent variables and with one single latent variable were also estimated.

The confirmatory factor analysis model (see Figure 1) showed that all items had high standardized coefficients and low standardized uniquenesses. The correlation between the *urge to move* and *pleasure* scales was high at $r = .78$. The target model had a very good fit: CFI = .978 and RMSEA = 0.055 (90% CI: 0.023, 0.088). Both the nested models with only one latent variable, $\chi^2(1) = 394, p < .001$, and with two uncorrelated latent variables, $\chi^2(1) = 6955, p < .001$, had a significantly worse fit compared to the target model. The reliability of the measurement models was estimated at $\alpha = .90$ (Cronbach's α) for the *urge to move*, and $\alpha = .97$ for *pleasure*.

Figure 2 shows a scatterplot of the *urge to move* ratings (based on items 1, 2, and 4) on the vertical axis and *pleasure* ratings (based on items 11, 13, and 15) on the horizontal axis (small symbols). The mean ratings for each stimulus are represented by the large symbols. Our predictions from Table 3 are more or less confirmed: "Superstition" (1) and "Bala" (2) were rated highly in both dimensions, whereas "Machine Gun" (7) and "My Pal Foot Foot" (8) obtained low ratings on both scales. "Sunrise" (5) and the Mahler passage (6) obtained high *pleasure* and low *urge to move* ratings. However, the mean ratings of "Tchip Tchip" (3) and "Hamdouchi" (4) occupy a central position, not the expected position in the upper left quadrant.

CONCLUSIONS

The results of Experiment 1 indicate that items 1, 2, and 4 are strong candidate items to measure participants' experience of an *urge to move* while listening to music. Similarly, items 11, 13, and 15 are strong candidates for the *pleasure* scale. The data for these six items had a good fit with the target model, and the six items measured the two constructs with high reliability. Experiment 1 also confirmed most of our expectations regarding the *urge to move* and *pleasure* qualities of the eight stimuli.

Experiment 2: Initial Validation of the Six-Item Questionnaire

A second listening experiment was carried out in order to verify the factor structure of the six items identified in Experiment 1 and to offer some evidence on the validity of the questionnaire. In the first experiment, the sample of participants was relatively small ($n = 56$), predominantly male, and a majority of participants reported to be either professional or amateur musicians. For Experiment 2, an independent and larger sample of participants was recruited that better represented the general population with respect to musical expertise and gender.
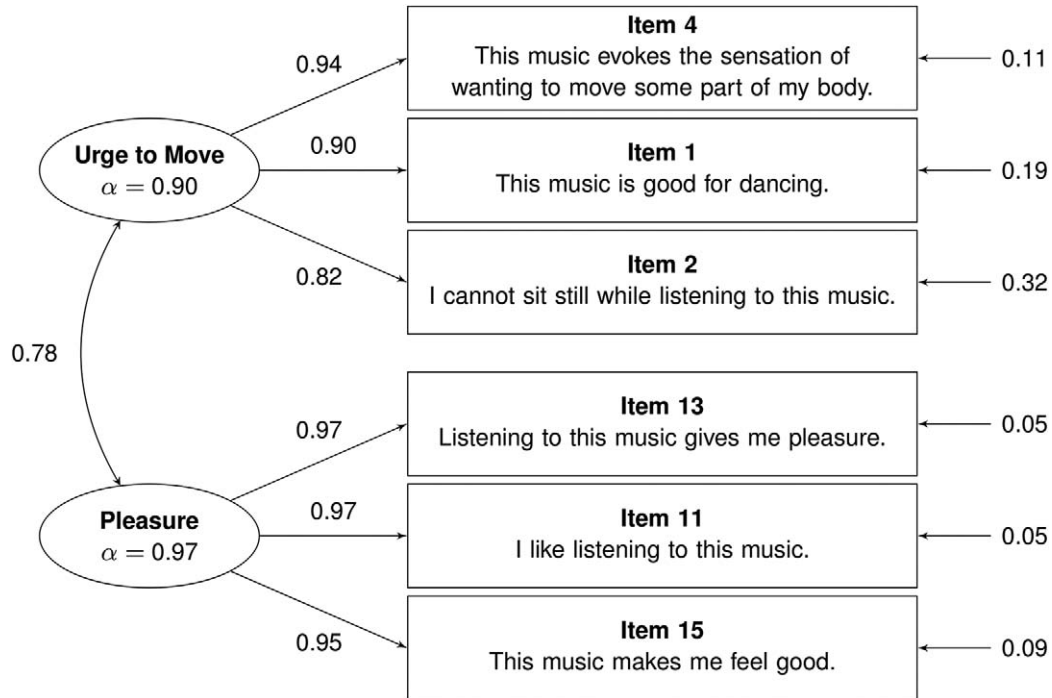
FIGURE 1. Confirmatory factor analysis of the target model with three selected indicators for each of the two constructs, *urge to move* and *pleasure* (Experiment 1). Items 4, 1, and 2 were selected to measure the *urge to move* construct. Items 13, 11, and 15 measure the *pleasure* construct (items are listed with increasing uniqueness for each construct). Double arrow: Pearson correlation coefficient; α: Cronbach's alpha; arrows from factors to items: factor coefficients; arrows from the right: uniquenesses.
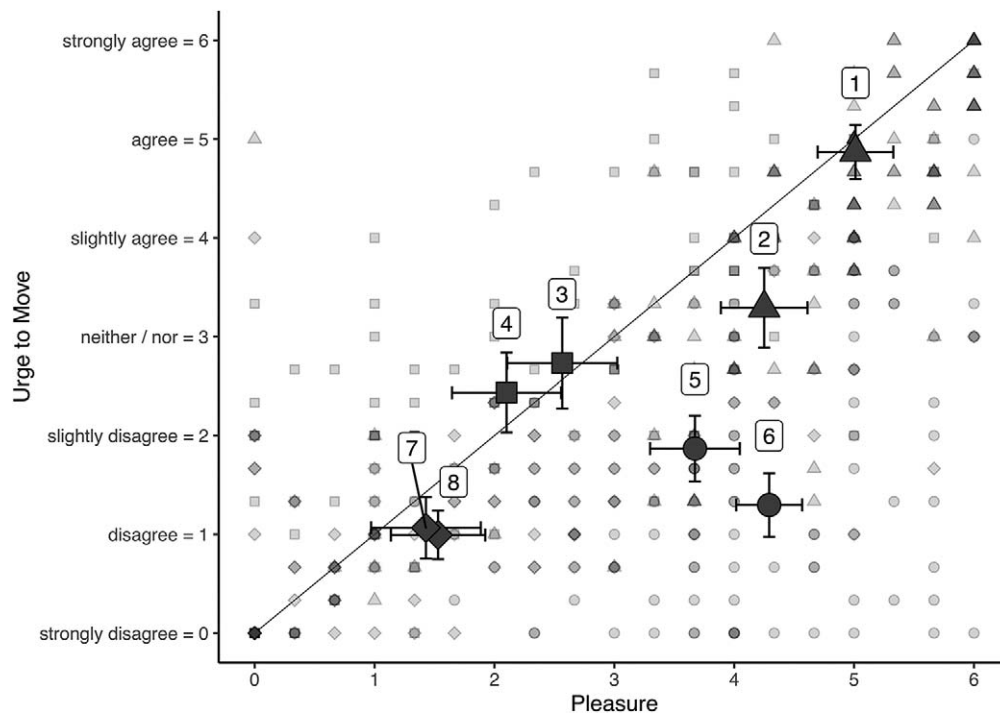
FIGURE 2. Scatterplot of *urge to move* ratings (vertical axis, mean of items 1, 2, and 4) against *pleasure* ratings (horizontal axis, mean of items 11, 13, and 15) for Experiment 1. Small semi-transparent symbols represent ratings on single observations. Large opaque symbols represent mean ratings per stimulus. The stimuli are numbered as follows: Superstition (1), Bala (2), Tchip Tchip (3), Hamdouchi (4), Sunrise (5), Mahler (6), Machine Gun (7), My Pal Foot Foot (8), see also Table 3. Error bars are 95% confidence intervals for the mean. The main diagonal is indicated as a grey line.
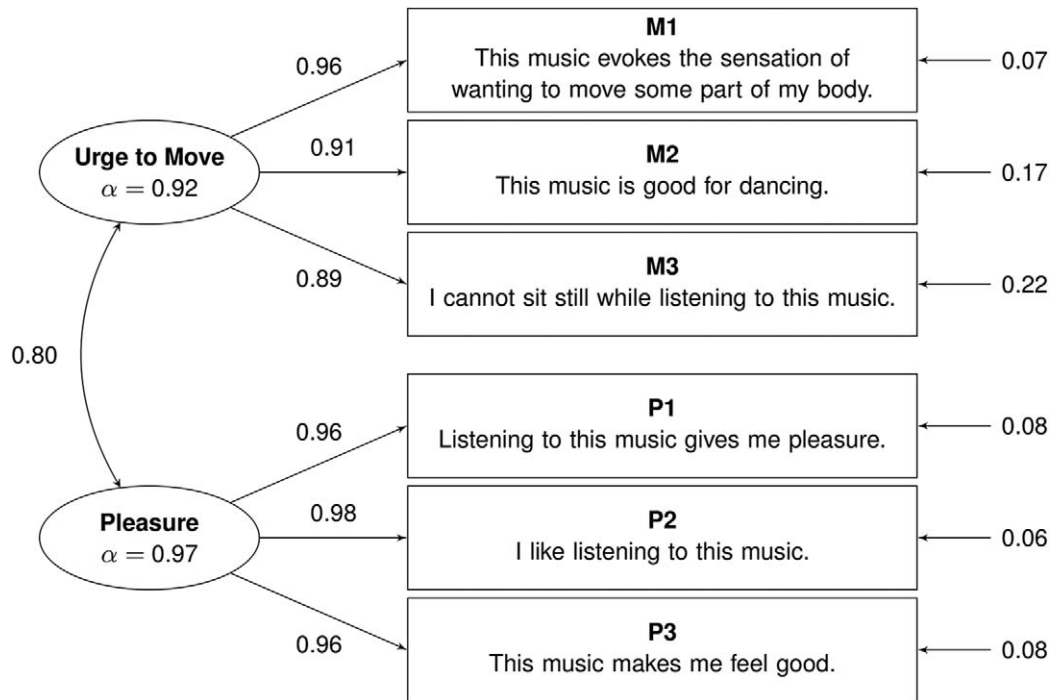
FIGURE 3. Confirmatory factor analysis of the target model with three indicators for each of the two constructs, *urge to move* and *pleasure* (Experiment 2). M1-M3 represent the items pertaining to the *urge to move* construct. P1-P3 represent items related to the *pleasure* construct. Double arrow: Pearson correlation coefficient; α: Cronbach's alpha; arrows from factors to items: factor coefficients; arrows from the right: uniquenesses.

METHOD

Procedure and stimuli. Participants were recruited through *Amazon's MTurk* platform. *MTurk* acts as an intermediary that connects providers of online work tasks with people who complete these tasks for a remuneration. We expected that participants recruited through *MTurk* would be balanced in terms of gender, and that only a low proportion of participants would have high musical expertise. Participants obtained an access code to carry out the listening experiment on the *SosciSurvey* platform. The procedure was identical to Experiment 1, but participants responded to the music using only the six selected questionnaire items (Figure 3). The experimental stimuli were the same as in Experiment 1 (Table 3) though shortened to the first 30 seconds each, because the questionnaire was much more compact in Experiment 2 (6 items) compared to Experiment 1 (22 items). After successfully finishing the survey (mean duration 13 minutes), participants obtained another access code to collect a remuneration of USD 2.50 on *MTurk*.

Participants. Of the 200 participants that finished the survey, three were excluded, because they stayed less

than 30 seconds on each page of the survey and therefore did not listen to the stimuli in their entirety. We screened the rating patterns of participants in order to detect anomalies (regular geometric rating patterns, invariant ratings) that would indicate a lack of motivation in participants. No participants were excluded due to their rating pattern.

Therefore, 197 participants (84 female) provided a complete set of responses. Of these participants, two self-identified as professional musicians, 37 as amateur musicians, 156 as music listeners, and two reported not to be interested in music. Participants had a mean age of 40 years (ranging from 24 to 73). This sample more plausibly represents a general population in terms of gender, musical expertise, and age, compared to the sample of Experiment 1.

Participants lived in the United States ($n = 158$), India ($n = 35$), and Italy ($n = 1$); three did not provide information on their country of residence. Therefore, geographically and culturally, the sample predominantly represented two different populations, namely residents of the United States and India. For comparison, the sample of Experiment 1 was mostly European.
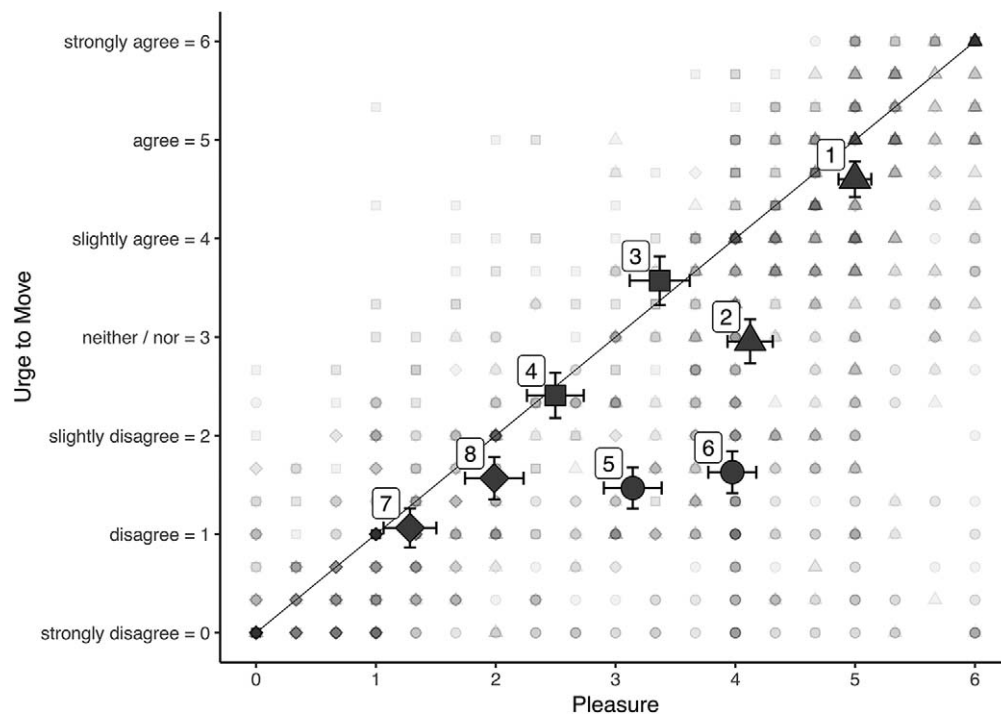
FIGURE 4. Scatterplot of *urge to move* (vertical axis) against *pleasure* (horizontal axis) ratings (Experiment 2). Small semi-transparent symbols represent ratings on single observations. Large opaque symbols represent mean ratings per stimulus. The stimuli are numbered as follows: Superstition (1), Bala (2), Tchip Tchip (3), Hamdouchi (4), Sunrise (5), Mahler (6), Machine Gun (7), My Pal Foot Foot (8), see also Table 3. Error bars are 95% confidence intervals for the mean. The main diagonal is indicated as a grey line.

Analysis. With eight experimental stimuli and 197 participants, a dataset of 1,576 complete observations was obtained. It was analysed using the *Lavaan* package's confirmatory factor analysis (CFA) function in *R*. Similar to Experiment 1, the target model consisted of two correlated factors representing the concepts *urge to move* and *pleasure*, each measured with three indicators.

RESULTS

Figure 3 shows the CFA factor structure of the target model calculated from the data collected in Experiment 2. Note that the questionnaire items have been relabeled M1–M3 and P1–P3 (whereby "M" stands for the *urge to move* and "P" for *pleasure*).

The target model assumed two correlated factors. It had a very good fit with CFI = .998 and RMSEA = 0.067 (90% CI: 0.052, 0.082). The fit of the target model was significantly better than the fit of nested alternative models, namely a model with only one latent factor, $\chi^2(1) = 1531, p < .001$, and a model with two orthogonal factors, $\chi^2(1) = 33390, p < .001$.

Both the *urge to move* (Cronbach's $\alpha = .92$) and *pleasure* ($\alpha = .97$) factors showed very high reliabilities. Each standardized factor coefficient was high, and the two factors were strongly correlated with each other at $r = .80$. The data from listening Experiment 2 confirms the factor structure proposed in Experiment 1. This corroborates the construct validity of the questionnaire: the *urge to move* and *pleasure* scales relate to the underlying constructs in a consistent way (Bollen, 1989, p. 188).

For each observation, an *urge to move* factor score was calculated as the mean of items M1, M2 and M3. Similarly, a *pleasure* factor score was calculated as the mean of P1 through P3.

Figure 4 shows a scatterplot of *urge to move* factor scores on the vertical axis and *pleasure* factor scores on the horizontal axis. Single datapoints are represented with small, semi-transparent symbols. Superimposed as bigger symbols are the mean factor scores for each of the eight stimuli with error bars (95% confidence intervals for the mean).

DISCUSSION

Figure 4 indicates that authors' expectations about how the stimuli would be rated were mostly confirmed (see Table 3, above), and they were consistent across both experiments (see also Figure 2):

→ Superstition (1) and Bala (2) were categorized as high *urge to move*/high *pleasure* stimuli by the researchers prior to the experiment, and accordingly obtained relatively high mean values on both scales.

→ Our prediction was also confirmed for Machine Gun (7) and My Pal Foot Foot (8) which both obtained low mean ratings on both scales.

→ Sunrise (5) and the slow Mahler movement (6) had been chosen to obtain low *urge to move* scores, which was confirmed, and high *pleasure* scores, which were more centralized (respectively lower) than we expected.

→ Finally, Tchip Tchip (3) and Hamdouchi (4) obtained high *urge to move* ratings as expected, but they also were rated relatively high on *pleasure*, which was against our prediction. In consequence, the mean ratings of Tchip Tchip (3) and Hamdouchi (4) occupy a central position in Figure 4 instead of a location in the upper left quadrant. Particularly, our selection of Tchip Tchip (3) as a representant of the low pleasure category seems to have been misguided: this music seems to trigger considerable pleasure in the surveyed population.

Stimuli from Western popular and classical music styles (Superstition, Mahler, Tchip Tchip) had a tendency to obtain higher ratings on both scales than stimuli from non-Western styles (Bala, Sunrise, Hamdouchi). This might be related to the fact that most respondents were from the United States; and they were more likely to be familiar with Western music compared to non-Western music. Previous research has shown that familiarity with a repertoire is positively associated with groove ratings (Senn, Bechtold, Hoesl, & Kilchenmann, 2019).

On the level of the single observation (small semi-transparent symbols in Figure 4), the *urge to move* ratings rarely exceeded the corresponding *pleasure* ratings. These data points (17.7% of all ratings) are located above the main diagonal shown in Figure 4. For some observations, the *urge to move* rating was equal to the *pleasure* rating (27.5% of the ratings). These data points are located on the main diagonal. In the majority of all cases, *pleasure* ratings were higher than the corresponding *urge to move* ratings (54.8%); these data points appear below the main diagonal.

The described constellation suggests that while *pleasure* might have been a precondition for the *urge to move*, the inverse was not necessarily true. Potentially, we are unlikely to feel an urge to dance, if we do not enjoy listening to the music. Yet, we may enjoy listening to music, even though we do not feel an urge to move. The Mahler stimulus seems to trigger this kind of reaction most strongly among the eight musical excerpts presented in this study. This interpretation of the relationship of *pleasure* and *urge to move* agrees with Matthews et al. (2019) and Senn, Rose, et al. (2019), who suggest that pleasure is a mediator for the urge to move. However, further corroboration is necessary, as this observation might be an artefact of a failure to select stimuli that truly inhabit the high *urge to move*/low *pleasure* quadrant of Figure 4 (top left).

The reliability estimates for the *urge to move* and *pleasure* scales are both very high. Survey length might be shortened in future applications of the questionnaire by removing one or even two items from a scale. When one of the three items was removed from the *urge to move* scale, reliability was estimated at Cronbach's $\alpha = .87$ (M1 removed), .89 (M2 removed), and .90 (M3 removed), respectively. When two items were removed, the Spearman-Brown prophecy formula (Brown, 1910) predicted a reliability of $\alpha = .79$ for the remaining single-item *urge to move* scale. Judging from the pattern coefficient values (Figure 3), the M1 single-item scale can be expected to be most closely aligned with the underlying *urge to move* construct and thus be the most reliable of the three scales.

When one of the three items was omitted from the *pleasure* scale, reliability decreased only slightly to $\alpha = .96$ (P1 removed), .95 (P2 removed), or .96 (P3 removed), respectively. When two items were removed, the Spearman-Brown prophecy formula still predicted an excellent reliability of $\alpha = .92$. P1 is likely to be the best single-item questionnaire for the *pleasure* scale, since it addresses the concept of pleasure explicitly.

Confirmatory factor analysis models were fitted separately to the data of the participants from the United States ($n = 158$) and India ($n = 35$) in order to investigate whether we need to expect strongly divergent responses when participants come from different cultural backgrounds. In the sub-sample from India, the reliability of the *urge to move* scale was slightly lower ($\alpha = .87$) than in the sub-sample from the U.S. ($\alpha = .93$). This was due to the fact that the responses on item M3 ("I cannot sit still while listening to this music") of participants from India were less well aligned with the *urge to move* scale (pattern coefficient = 0.68) than responses given by participants from the United States (pattern coefficient = 0.92). It is unclear why item M3 triggered different responses in participants from the two subsamples. Nevertheless, the reliability of the *urge to move* scale remained fairly high in the data provided by

the participants from India. We estimate that the small discrepancy does not disqualify the questionnaire from being used in a cross-cultural context.

Reducing the scales to single items (for example M1 for the *urge to move* and P1 for *pleasure*) is defensible, given that the items have good content validity and excellent (in the case of *pleasure*) or at least acceptable reliability (in the case of the *urge to move*). Nevertheless, we recommend using the three-item scales whenever it is reasonable to ask listeners to give three responses instead of only one. There are several reasons for this recommendation: First, there may be unexpected interactions between single items, populations, and stimuli (the differences in the M3 ratings by participants from India and the United States illustrate this point). Using multi-item scales dampens erratic behavior, and researchers will notice a drop in reliability when one item behaves unexpectedly. There is no such safeguard when single-item scales are used. Second, the composite three-item scales allow for 19 different rating outcomes per participant, whereas a single-item 7-point Likert scale only offers seven different outcomes. Consequently, the three-item scale is a more differentiated measure, and the data it produces are more likely to have favourable statistical properties. Third, groove models will grow more complicated in the near future, and they will move beyond regression models that simply estimate the direct effects of independent variables on groove response variables. Senn, Rose, et al. (2019) drafted a groove model that places the *urge to move* within a greater context of mental and bodily processes. These kinds of models can be studied using a structural equation modelling framework. Multi-item scales are a necessity in this context, because they provide the degrees of freedom required to fit the models. For these reasons, we recommend the use of the three-item scales whenever it is feasible and stimuli are rated as a whole.

However, the three-item scales will not be practical in a context where the *urge to move* or listening *pleasure* are studied as quantities that change in real time during the listening process. In this situation, single-item scales (M1 or P1) might be used instead with a continuous slider that is anchored by the seven response categories of the 7-point Likert scale (ranging from 0 = *strongly disagree* to 6 = *strongly agree*).

Conclusions

In music psychology the groove experience is defined as a pleasurable urge to move in response to music (Janata et al., 2012, p. 56). The aim of this study was to develop the *Experience of Groove Questionnaire* with an *urge to*

move and a *pleasure* scale, based on the Janata et al. (2012) definition. The score sheet of the final questionnaire can be found in the Appendix.

The development was carried out in three phases. In the first phase (Pre-Study), we generated 25 candidate questionnaire items and let groove research experts establish the items' content validity. In the second phase (Experiment 1), we studied the factor structure of these items and reduced the questionnaire to six items that were highly correlated with the two underlying *urge to move* and *pleasure* constructs (three items each). In the third phase (Experiment 2) we provided evidence for the construct validity and the reliability of the *urge to move* and *pleasure* scales in a listening experiment with an independent sample of participants.

Factor structures of the confirmatory factor analyses were very similar across Experiments 1 and 2, even though the samples of the two experiments differed strongly in terms of gender distribution, musical expertise, and cultural background (Europe, United States, India). We formulated the questionnaire items in compliance with Ogden's Basic English in order to make the questionnaire accessible to non-native English speakers. This suggests that the questionnaire is applicable to a general population. However, further research could translate the groove questionnaire to other languages to increase its accessibility.

The two scales of the groove questionnaire will prove to be useful for future studies in groove research and/or, more generally, in music psychology. They reliably measure the intensity of a music listener's inner *urge to move* and/or experienced *pleasure*. The scales assess these intensities from the listener's subjective point of view. They are compatible with the majority of groove definitions in music psychology. Due to the high reliability of the three-item scales, the number of items in either scale can be reduced while retaining acceptable reliability. If seven-point Likert scales (ranging from 0 to 6) are applied and factor scores are calculated as the mean of the subscales, results will be comparable across studies.

The two scales proved to be strongly and positively correlated. This dependency may be interpreted as follows: *pleasure* might be seen as a necessary precondition for the *urge to move*. The underlying hypothesis is that listeners who do not enjoy listening to the music are unlikely to feel an urge to move. Conversely, the *urge to move* might not be necessary for a listener in order to experience *pleasure*. This is a tentative interpretation that is compatible with the results of both experiments, but which could be a consequence of this study's stimuli selection. We welcome future studies that seek to expand the range of stimuli in order to further validate

the *Experience of Groove Questionnaire*, to explore its psychometric properties, and to investigate the relationship between pleasure and the urge to move.

Measuring the intensity of a person's *urge to move* and *pleasure* using a questionnaire has the limitations that are common for self-report psychometric instruments: the scales do not measure the intensity of the experience itself, but listeners' subjective impression of intensity. Consequently, measurements will be influenced by psychological biases that are frequently observed in self-report questionnaires. The effects of these biases will remain largely undiscovered until other measurement methods are used simultaneously to assess the urge to move or pleasure. Potentially, the groove questionnaire can be further validated in the future by correlating the ratings with objective measurements of participants' spontaneous body movement (e.g. using motion capture technology), or with neural or hormonal signatures related to the experience of an urge to move or pleasure.

The *Experience of Groove Questionnaire* has primarily been designed as a basic measurement tool for psychological groove research. The *urge to move* and *pleasure* scales can be used as response variables in studies that aim to model the intensity of listeners' experience of groove depending on musical, personal, situational or other factors (Senn, Rose, et al., 2019).

However, the questionnaire can also be applied to evaluate music/stimuli in a variety of other research contexts. The two scales might prove to be useful in sports psychology. Athletes synchronize their movements to music in a variety of training settings in order to increase motivation and to counteract fatigue (Karageorghis & Priest, 2012). If music can be proven to increase athletes' *urge to move* and feeling of *pleasure*, it is likely that it also increases their motivation for training and has a positive effect on their endurance. The relationship between the *Experience of Groove* scales and the six-item single-scale *Brunel Music Rating Inventory* that measures the motivational effect of music in sports and exercise needs to be investigated (Clark, Baker, Peiris, Shoebridge, & Taylor, 2016; Karageorghis, 2020). Further, the *pleasure* scale measures a fundamental aesthetic response to music (Cupchik & Gebotys, 1990). It could contribute to an empirical investigation of aesthetic judgments on music in general. Finally, recent research reported that music facilitates sensorimotor synchronisation in people with Parkinson's (Rose, Delevoye-Turrell, Ott, Annett, & Lovatt, 2019). The *Experience of Groove Questionnaire* may help to identify music that triggers an *urge to move* in people with Parkinson's. Music that scores high on both scales among people with Parkinson's can potentially be successfully applied in music therapy and rehabilitation settings.

Author Note

Correspondence concerning this article should be addressed to Olivier Senn, Lucerne University of Applied Sciences and Arts, Music Department, Arsenalstrasse 28a, 6010 Kriens, Switzerland, olivier.senn@hslu.ch.

References

BEAUDUCEL, A., & HERZBERG, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

BERLINER, P. F. (1994). *Thinking in jazz: The infinite art of improvisation.* Chicago, IL: University of Chicago Press.

BOLLEN, K. A. (1989). *Structural equations with latent variables.* New York: J. Wiley and Sons.

BROWN, T. A. (2006). *Confirmatory factor analysis for applied research.* New York: Guilford Press.

BROWN, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 1904-1920, 3*(3), 296–322. https://doi.org/10.1111/j.2044-8295.1910.tb00207.x

BURGER, B., THOMPSON, M. R., LUCK, G., SAARIKALLIO, S., & TOIVIAINEN, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology, 4*(183). https://doi.org/10.3389/fpsyg.2013.00183

BUTTERFIELD, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum, 33*, 3–26. https://doi.org/10.1525/mts.2011.33.1.3

CÂMARA, G. S., & DANIELSEN, A. (2018). Groove. *The Oxford handbook of critical concepts in music theory.* https://doi.org/10.1093/oxfordhb/9780190454746.013.17

CAMERON, D. J., ZIOGA, I., LINDSEN, J. P., PEARCE, M. T., WIGGINS, G. A., POTTER, K., & BHATTACHARYA, J. (2019). Neural entrainment is associated with subjective groove and complexity for performed but not mechanical musical rhythms. *Experimental Brain Research, 237,* 1981–1991. https://doi.org/10.1007/s00221-019-05557-4

CHUSID, I. (2000). *Songs in the key of z: The curious universe of outsider music.* Chicago, IL: Chicago Review Press.

CLARK, I. N., BAKER, F. A., PEIRIS, C. L., SHOEBRIDGE, G., & TAYLOR, N. F. (2016). The Brunel Music Rating Inventory-2 is a reliable and valid instrument for older cardiac rehabilitation patients selecting music for exercise. *Psychology of Music, 44,* 249–262. https://doi.org/10.1177/0305735614565830

CUPCHIK, G. C., & GEBOTYS, R. J. (1990). Interest and pleasure as dimensions of aesthetic response. *Empirical Studies of the Arts, 8,* 1–14. https://doi.org/10.2190/L789-TPPY-BD2Q-T7TW

DANIELSEN, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament.* Middletown, CT: Wesleyan University Press.

DANIELSEN, A., HAUGEN, M. R., & JENSENIUS, A. R. (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. *Timing and Time Perception, 3,* 133–154. http://dx.doi.org/10.1163/22134468-00002043

DAVIES, M., MADISON, G., SILVA, P., & GOUYON, F. (2013). The effect of microtiming deviations on the perception of groove in short rhythms. *Music Perception, 30,* 497–510. https://doi.org/10.1525/mp.2013.30.5.497

DIAMANTOPOULOS, A., SARSTEDT, M., FUCHS, C., WILCZYNSKI, P., & KAISER, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science, 40,* 434–449. https://doi.org/10.1007/s11747-011-0300-3

DISTEFANO, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 9,* 327–346. https://doi.org/10.1207/S15328007SEM0903_2

DOFFMAN, M. R. (2008). *Feeling the groove: Shared time and its meanings for three jazz trios* (PhD thesis). The Open University, Milton Keynes, England.

DOLAN, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology, 47,* 309–326. https://doi.org/10.1111/j.2044-8317.1994.tb01039.x

EAVES, D. L., GRIFFITHS, N., BURRIDGE, E., MCBAIN, T., & BUTCHER, N. (2019). Seeing a drummer's performance modulates the subjective experience of groove while listening to popular music drum patterns. *Musicae Scientiae.* https://doi.org/10.1177/1029864919825776

ETANI, T., MARUI, A., KAWASE, S., & KELLER, P. (2018). Optimal tempo for groove: Its relation to directions of body movement and Japanese nori. *Frontiers in Psychology, 9*(462). https://doi.org/10.3389/fpsyg.2018.00462

FRÜHAUF, J., KOPIEZ, R., & PLATZ, F. (2013). Music on the timing grid: The influence of microtiming on the perceived groove quality of a simple drum pattern performance. *Musicae Scientiae, 17*(2), 246–260. https://doi.org/10.1177/1029864913486793

GORSUCH, R. L. (2015). *Factor analysis: Classic edition.* New York: Routledge.

HOSKEN, F. (2020). The subjective, human experience of groove: A phenomenological investigation. *Psychology of Music, 48,* 182–198 https://doi.org/10.1177/0305735618792440

HOVE, M. J., MARTINEZ, S. A., & STUPACHER, J. (2020). Feel the bass: Music presented to tactile and auditory modalities increases aesthetic appreciation and body movement. *Journal of Experimental Psychology: General, 149.* 1137–1147.

IYER, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception, 19,* 387–414.

JANATA, P., TOMIC, S. T., & HABERMAN, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General, 141,* 54–75. https://doi.org/10.1037/a0024208

KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23,* 187–200. https://doi.org/10.1007/BF02289233

KARAGEORGHIS, C. I. (2020). Music-related interventions in sport and exercise. In G. Tenenbaum & R. C. Eklund (Ed.), *Handbook of sport psychology* (4th edition, pp. 929–949). Hoboken, NJ: Wiley.

KARAGEORGHIS, C. I., & PRIEST, D.-L. (2012). Music in the exercise domain: A review and synthesis (Part II). *International Review of Sport and Exercise Psychology, 5,* 67–84. https://doi.org/10.1080/1750984X.2011.631027

KEIL, C. (2010). Defining "groove." *PopScriptum, 11.* http://www2.hu-berlin.de/fpm/popscrip/themen/pst11/pst11_keil02.html

KERNFELD, B. (2003). Groove. In *Grove music online.* London, UK: Macmillan Press.

KILCHENMANN, L., & SENN, O. (2015). Microtiming in Swing and Funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology: Performance Science, 6*(1232). https://doi.org/10.3389/fpsyg.2015.01232

LABBÉ, C., & GRANDJEAN, D. (2014). Musical emotions predicted by feelings of entrainment. *Music Perception, 32,* 170–185.

LUSTIG, E., & TAN, I. (2019). All about that bass: Audio filters on basslines determine groove and liking in electronic dance music. *Psychology of Music.* https://doi.org/10.1177/0305735619836275

MADISON, G. (2001). Different kinds of groove in jazz and dance music as indicated by listeners' ratings. *Proceedings of the VII International Symposium on Systematic and Comparative Musicology, III International Conference on Cognitive Musicology* (pp. 108–112). Jyväskylä, Finland: University of Jyväskylä.

MADISON, G. (2003). Perception of jazz and other groove-based music as a function of tempo. *Proceedings of the 5th Triennial ESCOM Conference* (pp. 365–367). Hannover, Germany: ESCOM.

MADISON, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception, 24,* 201–208. https://doi.org/10.1525/mp.2006.24.issue-2

MADISON, G., GOUYON, F., ULLÉN, F., & HÖRNSTRÖM, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology. Human Perception and Performance, 37,* 1578–1594. https://doi.org/10.1037/a0024323

MADISON, G., & SIOROS, G. (2014). What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology, 5*(894). https://doi.org/10.3389/fpsyg.2014.00894

MATSUSHITA, S., & NOMURA, S. (2016). The asymmetrical influence of timing asynchrony of bass guitar and drum sounds on groove. *Music Perception, 34,* 123–131. https://doi.org/10.1525/mp.2016.34.2.123

MATTHEWS, T. E., WITEK, M. A. G., HEGGLI, O. A., PENHUNE, V. B., & VUUST, P. (2019). The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PLOS ONE, 14,* e0204539. https://doi.org/10.1371/journal.pone.0204539

McCOACH, D. B., GABLE, R. K., & MADURA, J. P. (2013). *Instrument development in the affective domain: School and corporate applications* (3rd ed.). New York: Springer.

MONSON, I. (1996). *Saying something: Jazz improvisation and interaction.* Chicago, IL: University of Chicago Press.

NOZARADAN, S., PERETZ, I., MISSAL, M., & MOURAUX, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience, 31,* 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

NOZARADAN, S., SCHÖNWIESNER, M., KELLER, P. E., LENC, T., & LEHMANN, A. (2018). Neural bases of rhythmic entrainment in humans: Critical transformation between cortical and lower-level representations of auditory rhythm. *The European Journal of Neuroscience, 47,* 321–332. https://doi.org/10.1111/ejn.13826

O'CONNOR, R. (1996). The fifty most significant Indie records. In S. Schinder (Ed.), *Rolling Stone's alt rock-a-rama* (pp. 60–74). New York: Delta.

OGDEN, C. K. (1948). *The system of basic English.* New York: Harcourt, Brace and Co.

PETT, M. A., LACKEY, N. R., & SULLIVAN, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research.* Thousand Oaks, CA: SAGE.

PRESSING, J. (2002). Black atlantic rhythm: Its computational and transcultural foundations. *Music Perception, 19,* 285–310.

ROHOLT, T. C. (2014). *Groove: A phenomenology of rhythmic nuance.* New York: Bloomsbury Publishing USA.

ROSE, D., DELEVOYE-TURRELL, Y., OTT, L., ANNETT, L. E., & LOVATT, P. J. (2019). Music and metronomes differentially impact motor timing in people with and without Parkinson's disease: Effects of slow, medium, and fast tempi on entrainment and synchronization performances in finger tapping, toe tapping, and stepping on the spot tasks. *Parkinson's Disease, 2019*(6530838), 1–18. https://doi.org/10.1155/2019/6530838

SENN, O., BECHTOLD, T. A., HOESL, F., & KILCHENMANN, L. (2019). Taste and familiarity affect the experience of groove in popular music. *Musicae Scientiae,* 1–22. https://doi.org/10.1177/1029864919839172

SENN, O., KILCHENMANN, L., BECHTOLD, T., & HOESL, F. (2018). Groove in drum patterns as a function of both rhythmic properties and listeners' attitudes. *PLOS ONE, 13*(6), e0199604. https://doi.org/10.1371/journal.pone.0199604

SENN, O., KILCHENMANN, L., VON GEORGI, R., & BULLERJAHN, C. (2016). The effect of expert performance microtiming on listeners' experience of groove in swing or funk music. *Frontiers in Psychology: Performance Science, 7*(1487). https://doi.org/10.3389/fpsyg.2016.01487

SENN, O., ROSE, D., BECHTOLD, T., KILCHENMANN, L., HOESL, F., JERJEN, R., BALDASSARRE, A., & ALESSANDRI, E. (2019). Preliminaries to a psychological model of musical groove. *Frontiers in Psychology, 10*(1228). https://doi.org/10.3389/fpsyg.2019.01228

SIOROS, G., MIRON, M., DAVIES, M., GOUYON, F., & MADISON, G. (2014). Syncopation creates the sensation of groove in synthesized music examples. *Frontiers in Psychology, 5*(1036). https://doi.org/10.3389/fpsyg.2014.01036

STUPACHER, J., HOVE, M. J., & JANATA, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception, 33,* 571–589. https://doi.org/10.1525/mp.2016.33.5.571

STUPACHER, J., HOVE, M. J., NOVEMBRE, G., SCHÜTZ-BOSBACH, S., & KELLER, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition, 82,* 127–136. https://doi.org/10.1016/j.bandc.2013.03.003

STUPACHER, J., WITTE, M., HOVE, M. J., & WOOD, G. (2016). Neural entrainment in drum rhythms with silent breaks: Evidence from steady-state evoked and event-related potentials. *Journal of Cognitive Neuroscience*, 28, 1865–1877. https://doi.org/10.1162/jocn_a_01013

SWARBRICK, D., BOSNYAK, D., LIVINGSTONE, S. R., BANSAL, J., MARSH-ROLLO, S., WOOLHOUSE, M. H., & TRAINOR, L. J. (2019). How live music moves us: Head movement differences in audiences to live versus recorded music. *Frontiers in Psychology*, 9(2682). https://doi.org/10.3389/fpsyg.2018.02682

TEMPLER, B. (2007). Less is more: Reconfiguring simplified models for English as a lingua franca. *The New English Teacher*, 1(2), 11–25.

THOMPSON, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington D.C.: American Psychological Association.

VON GEORGI, R., BULLERJAHN, C., SENN, O., & KILCHENMANN, L. (2016). *Emotional assessment of groove*. Retrieved from https://www.researchgate.net/publication/308202805_Emotional_Assessment_of_Groove

WESOLOWSKI, B. C., & HOFMANN, A. (2016). There's more to groove than bass in electronic dance music: Why some people won't dance to techno. *PLOS ONE*, 11(10), e0163938. https://doi.org/10.1371/journal.pone.0163938

WITEK, M. A. G. (2016). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, 36, 138–160. https://doi.org/10.1111/musa.12082

WITEK, M. A. G., CLARKE, E. F., WALLENTIN, M., KRINGELBACH, M. L., & VUUST, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, 9(4). https://doi.org/10.1371/journal.pone.0094446

ZBIKOWSKI, L. M. (2004). Modelling the groove: Conceptual structure and popular music. *Journal of the Royal Musical Association*, 129, 272–297.

ZENTNER, M., GRANDJEAN, D., & SCHERER, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8, 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Appendix

TABLE A. *Questionnaire Items Used in Past Groove Research to Measure the Strength of Music Listeners'* Urge to Move *in Response to Music.*

Source	Item
Madison, 2001, p. 2	Makes me want to move some part of my body
Madison, 2003, p. 2	Makes me want to move some part of my body
Madison, 2006, p. 201	Wanting to move some part of the body in relation to some aspect of the sound pattern
Madison et al., 2011, p. 1580	Evokes the sensation of wanting to move some part of the body
Janata et al., 2012, p. 58	Rate the degree to which you feel the music you are hearing 'grooves'.
Davies et al., 2013, p. 497	Wanting to move some part of the body in relation to some aspect of the sound pattern
Frühauf et al., 2013, p. 250	This version animates me to move my body.
Stupacher et al., 2013, p. 130	An aspect of music that makes you want to move
Madison & Sioros, 2014, p. 1	The sensation of wanting to move some part of your body in relation to some aspect of the music
Sioros et al., 2014, p. 4	The sensation of wanting to move some part of your body in relation to some aspect of the music
Witek et al., 2014, p. 4	To what extent does this rhythm make you want to move?
Labbé & Grandjean, 2014, p. 173	To what extent did you feel like dancing?
	To what extent did you feel like moving?
	To what extent did you feel energized?
	To what extent did you feel the rhythm in your body?
	To what extent did you feel animated?
Matsushita & Nomura, 2016, p. 125	Feeling of drive to move your body parts along with the music
Senn et al., 2016, p. 6	I had the impression that my head moved with the rhythm.
	I felt like tapping my foot with the music.
	The music stimulated me to bop along.
	The music stimulated me to clap along or click my fingers.
Stupacher, Hove, & Janata, 2016, p. 579	To what extent did you feel that the musical excerpt grooved?
	To what extent did you feel "in the groove" while listening to the excerpt?
Wesolowski & Hofmann, 2016, p. 7	This music makes me want to tap my feet.
	This music makes me want to nod my head.
	This musical sample makes me want to dance.

TABLE A. *(continued)*

Source	Item
Etani et al., 2018, p. 5	I want to move my body.
	I want to move my body back and forth.
	I want to move my body side to side.
	I feel like dancing.
	I feel like my body resonates with the rhythm.
Senn et al., 2018, p. 11	I would like to dance to this music.
	This music is great for a party.
Cameron et al., 2019, p. 5	How much did this rhythm make you want to move?
Eaves et al., 2019, p. 8	To what extent did you feel the audio made you want to move?
Lustig & Tan, 2019, p. 5	How much does this music make you want to move?
Matthews et al., 2019, p. 6	How much does this musical pattern make you want to move?

TABLE B. *Questionnaire Items Used in Past Groove Research to Measure the Strength of Music Listeners' Experience of Pleasure.*

Source	Item
Madison, 2003, p. 2	I like it.
Madison et al., 2011, p. 1580	You like the music and wish to continue listening.
Frühauf et al., 2013, p. 250	I like this version.
Witek et al., 2014, p. 4	How much pleasure do you experience listening to this rhythm?
Matsushita & Nomura, 2016, p. 125	Feeling of pleasantness because of the music
Stupacher, Hove, & Janata, 2016, p. 579	How much did you enjoy what was just played?
Etani et al., 2018, p. 5	I feel pleasure.
Senn et al., 2018, p. 11	I like to listen to this music.
Cameron et al. 2019, p. 5	How much pleasure do you experience listening to this rhythm?
Lustig & Tan, 2019, p. 5	How much do you like this music?
Matthews et al., 2019, p. 6	How much pleasure do you experience listening to this musical pattern?

Discography

Stimulus	Track	Starting Time
1	Stevie Wonder (1972). *Superstition*. On *Talking Book*. Detroit, USA: Tamla.	01:24
2	Bassekou Kouyate & Ngoni Ba (2007). *Bala*. On *Segu Blue*. München, Germany: Out Here Records.	00:13
3	Cash & Carry (1973). *Tchip Tchip (Single)*. On *Tchip Tchip*. Netherlands: CNR.	00:00
4	Maleem Mahmoud Ghania with Pharoah Sanders (1994). *Hamdouchi*. On *The Trance Of Seven Colors*. New York City, USA: Axiom.	03:37
5	The Bulgarian Voices Angelite & Moscow Art Trio With Huun-Huur-Tu (1998). *Sunrise*. On *Mountain Tale*. Bremen, Germany: JARO Medien GmbH.	00:38
6	New York Philharmonic Orchestra & Leonard Bernstein (1989). *Symphony No. 3 Movement VI*. On *The Complete Mahler Symphonies*. Berlin, Germany: Sony Classical.	08:25
7	The Peter Brötzmann Octet (1986). *Machine Gun (Second Take)*. On *Machine Gun*. Germany: BRÖ.	11:07
8	The Shaggs (1969). *My Pal Foot Foot*. On *Philosophy Of The World*. USA: Third World Recordings.	00:20

Experience of Groove Questionnaire: Score Sheet

Senn, O., Bechtold, T., Rose, D., Câmara, G. S., Düvel, N., Jerjen, R., Kilchenmann, L., Hoesl, F., Baldassarre, A., Alessandri, E. (2020). Experience of Groove Questionnaire: Instrument Development and Initial Validation. *Music Perception*, *38*, 46–65.

Urge to Move Scale

Item	Strongly disagree (0)	Disagree (1)	Slightly disagree (2)	Neither agree nor disagree (3)	Slightly agree (4)	Agree (5)	Strongly agree (6)
M1 This music evokes the sensation of wanting to move some part of my body.	☐	☐	☐	☐	☐	☐	☐
M2 This music is good for dancing.	☐	☐	☐	☐	☐	☐	☐
M3 I cannot sit still while listening to this music.	☐	☐	☐	☐	☐	☐	☐

Pleasure Scale

Item	Strongly disagree (0)	Disagree (1)	Slightly disagree (2)	Neither agree nor disagree (3)	Slightly agree (4)	Agree (5)	Strongly agree (6)
P1 Listening to this music gives me pleasure.	☐	☐	☐	☐	☐	☐	☐
P2 I like listening to this music.	☐	☐	☐	☐	☐	☐	☐
P3 This music makes me feel good.	☐	☐	☐	☐	☐	☐	☐