Manuel Anglada-Tort, Heather Thueringer, Diana Omigie

The Busking Experiment: A Field Study Measuring Behavioural Responses to Street Music Performances

Anglada-Tort, M., Thueringer, H., & Omigie, D. (2019). The busking experiment: A field study measuring behavioral responses to street music performances. Psychomusicology: Music, Mind, and Brain, 29(1), 46–55. https://doi.org/10.1037/pmu0000236

The Busking Experiment: A Field Study Measuring Behavioural Responses to Street Music Performances

Manuel Anglada-Tort, Heather Thueringer, and Diana Omigie

Audio Communication, Technische Universität Berlin, Berlin, Germany

Department of Psychology, Goldsmiths, University of London, London, UK

Author Note

Manuel Anglada-Tort, Department of Audio Communication, Technische Universität Berlin, Berlin, Germany; Heather Thueringer, Department of Psychology, Goldsmiths, University of London, London, United Kingdom; Diana Omigie, Department of Psychology, Goldsmiths, University of London, London, United Kingdom. This work was based on the dissertation project of the MSc in Music, Mind, and Brain, Goldsmiths, University of London, by Heather Thueringer. The work and dataset have not previously been disseminated elsewhere.

Acknowledgements. We would like to thank Sonia Cassius-Clay for the helpful first-hand insight into busking as well as performing in the 24 experimental sessions conducted in this study. This work was supported by a PhD studentship from the “Studienstiftung des Deutschen Volkes” (Bonn, Germany) awarded to Manuel Anglada-Tort.

Correspondence concerning this paper should be addressed to Manuel Anglada-Tort, Department of Audio Communication, Technische Universität Berlin, Berlin, Germany. E-mail: m.angladatort@campus.tu-berlin.de

Abstract

A field experiment was conducted with a professional busker in the London Underground over the course of 24 days. Its aim was to investigate the extent to which performative aspects influence behavioural responses to music street performances. Two aspects of the performance were manipulated: familiarity of the music (familiar vs. unfamiliar) and body movements (expressive vs. restricted). The amount of money donated and number of people who donated were recorded. A total of 278 people donated over the experiment. The music stimuli, which was selected in an online study to differ only in familiarity, had been previously recorded by the busker. During the experimental sessions, the busker lip-synced to the pre-recorded recordings. Thus, the audio input in the experiment remained identical across sessions and the only variables that changed across conditions were the familiarity of the music and the expressivity of performed body movements. The results indicated that neither music familiarity nor performer's body movements had a significant impact on the amount of money donated (Rm^2= .033) nor the number of donors (Rm^2= .023). These results do not support previous literature on the influence of familiarity and performers' body movements, typically conducted in lab and artificial environments. The findings are further discussed with regard to potential extraneous variables that are crucial to control for (i.e., location of the performance, physical appearance, the bandwagon effect) and the advantages of field versus laboratory experiments. A novel research framework to study music judgements and behaviour is introduced, namely, the behavioural economics of music.

Keywords: busking, street performance, familiarity, body movements, field study

The Busking Experiment: A Field Study Measuring Behavioural Responses to Street Music Performances

"Busking is free, it's for everyone […] There's no smoke and mirrors – if people don't like it, they walk away" – Passenger (Foster, 2014)

Busking – or street performance for money – has been a popular practice in cities' public spaces for centuries (Cohen & Greenwood, 1981). As early as the 11th century, *troubadours* and *jongleurs* were entertaining the citizens of France, and in the 12th century, Germany was filled with *Minnesingers* and *Spielleute* (Smith, 1996). Since then, buskers have continued the tradition of street entertainment to the present day. However, despite the long history of street performance and the prevalence of buskers in most major cities across the globe, there has been remarkably little research conducted on this topic within the field of music psychology.

The majority of the literature on street musicians has focused on the history of busking (Campbell, 1981; Cohen & Greenwood, 1981; Smith, 1996) and single case studies about individual buskers, exploring the meaning and motivations behind busking practice (Jeffreys & Wang, 2012; Rebeiro Gruhl, 2017; Williams, 2016). Other studies have approached the topic of busking within the fields of economics (Kushner & Brooks, 2000), law (Quilter & McNamara, 2015; McNamara & Quilter, 2016), and ethnography as well as ethnomusicology (Breyley, 2016; Marina, 2018; Wong, 2016). However, none of these studies used a scientific approach to measure people's behavioural responses to street music performances or to explore potentially relevant factors mediating successful busking. Here, we take a behavioural economics approach, whereby monetary donations are monitored, in order to examine the extent to which two factors known to influence performance appreciation in the lab also have an influence in everyday settings. We therefore applied field research methodology commonly used by

behavioural scientists and experimental economists to investigate donating behaviour and charitable giving in the real world (e.g., Ebeling, Feldhaus, & Femdrich, 2017; Ekström, 2012; Khadjavi, 2016; Moussaoui, Naef, Tissot, & Desrichard, 2016; Olda & Ichihashi, 2016).

To the best of our knowledge, a study from Lemay and Bates (2013) is the only attempt in the scientific literature to investigate mediating factors contributing to busker donations. A sample of 103 undergraduate students were surveyed on their religion and attitudes toward busking. The best predictive model of giving to buskers was a three variable solution consisting of low religious fundamentalism, less experienced irritation toward buskers, and prior experience of giving to the homeless (Lemay & Bayes, 2013). Nevertheless, that study is limited in its reliance solely on survey methodology and a sample of undergraduate students, instead of measuring actual behaviour in real-world situations. Thus, a main motivation of the present study was to design a field experiment that investigates the impact of different performative aspects on people's behavioural responses to buskers; and in doing so, allowing the collection of raw data in a natural busking environment.

Two additional questions guided the current research, namely: What makes a successful street musician? And which aspects of the performative act might influence people's behavioural responses? To address these questions, we focused on two potential mediating factors that may be expected to influence the amount of donations and number of donors to busker performances. These were the familiarity of the music and the expressivity of the performer's body movements. The connection between familiarity and music enjoyment has been extensively investigated through the "mere exposure effect" (Zajonc, 1968), with most studies showing that liking for music increases with repeated exposure, or familiarity (see North & Hargreaves, 2008, for a

review). This effect has also been found in the evaluation of identical music performances (Anglada-Tort & Müllensifen, 2017; Korger & Margulis, 2016). Moreover, familiarity plays an important role in the emotional engagement of listeners with music (Pereira, Texeira, Figueiredo, Xavier, Castro, & Brattico, 2011); and familiar music has been positively associated with participants' willingness to pay for music (Tavani, Caroff, Storme, and Colange, 2016). Therefore, from a busker's point of view, the evidence appears overwhelmingly in favour of using familiar music stimuli over unfamiliar to create positive affect and, therefore, maximize profits.

As musicians often make expressive gestures and body movements while performing, the other performative aspect investigated in the present study was the expressivity of the busker's body movements. Previous studies suggest an influence of performer's body movements on the perception of emotion in music performances (Dahl & Friberg, 2004, 2007; Castellano, Mortillaro, Camurri, Volpe & Scherer, 2008; Chapados and Levitin, 2008; Timmers, Marolt, Camurri & Volpe, 2006; Vines, Krumhansl, Wanderley, & Levitin, 2006; Vines, Wanderley, Nuzzo, Levitin & Krumhansl, 2003). Furthermore, a meta-analysis of 15 studies revealed that the visual component of music performances plays an important role in their evaluation (i.e., liking, expressiveness, and overall quality), with a medium effect size on average (Platz & Kopiez, 2012). Thus, one could predict that buskers that express themselves non-verbally through body movements, may generate more profits than less expressive street musicians.

Field experiments offer important advantages compared to lab studies. However, field research is very scarce in the field of music psychology, where the majority of studies are conducted in experimental or lab settings (Hallam, Cross, & Thaut, 2016; some exceptions are Jacob, Guéguen, & Boulbry, 2010; North, Tarrant, & Hargreaves,

2004; Ruth, 2017). Controlled studies conducted in labs and other artificial environments are susceptible, amongst others, to two major problems (Carpenter, Harrison, & List, 2005; Reis & Judd, 2000): a lack of external validity – the extent to which the results are generalisable beyond the research setting and participants pool – and a lack of ecological validity – the degree to which the results apply to the real-world situation under study -. One can justify these problems by the high levels of internal validity - the extent to which an experiment controls for confounding variables – enabled by lab experiments. Nevertheless, it is also possible to control carefully for confounding variables in field research (Carpenter et al., 2005). The effects of familiarity and body movements on listeners' perception and appreciation of music have been well documented in lab settings (see North & Hargreaves, 2008; Platz & Kopiez, 2012, for reviews). Yet, are these findings reproducible outside of the lab and under real-world conditions? The current research addresses this question with the aid of a novel experimental design that carefully controls for potential confounding variables while enabling the measurement of people's economic responses to street music performances in a natural busking environment.

The present study aimed to investigate the extent to which music familiarity and expressivity of body movements influence behavioural responses to street music performances. A field experiment was conducted with a professional busker in the London Underground over 24 days. The amount of donations and number of donors were the measured dependent variables. Participants were London commuters and were not aware of taking part in a scientific study. Based on the literature outlined above, the following two hypotheses formed the bases for the current research:

i. As familiarity effects on music evaluation have been consistently demonstrated within a wide variety of experimental settings and stimuli

(Anglada-Tort & Müllensifen, 2017; Korger & Margulis, 2016; Pereira et al., 2011; Tavani et al., 2016; see North & Hargreaves, 2008, for a review), it was hypothesised that busking performances using familiar music will lead to a higher amount of donations and number of donors than performances using unfamiliar music.

ii. Based on previous research on the influence of performers' body movements influence on music performances (Dahl & Friberg, 2004, 2007; Castellano et al., 2008; Chapados and Levitin, 2008; Timmers et al., 2006; Vines et al., 2006; Vines et al., 2003), it was hypothesised that busking performances using expressive body movements will lead to a higher amount of donations and number of donors than performances using restricted body movements.

Due to the lack of published literature on the interaction between familiarity effects and body movements, no specific hypotheses are presented regarding the interaction term.

Method

Participants

Participants were commuters in the London Underground's Waterloo Station who happened to pass by during the music performances. Participants were unaware they were involved in research of any kind. Due to the location of the experiment, ethical considerations, and the nature of the study itself, cameras recording footage for the study did not capture faces of participants but only filmed the busker's donation bag and the feet of people walking nearby. The total number of people who passed within aural and visual range of the busker during the 24 sessions could not be estimated.

However, the total number of donations given over the experiment was 278. Note also that it is possible that these 278 donations may have included more than one donation from the same donor (who may have gone past the singer on more than one day).

Design

This research was granted ethical approval by the Ethics Committee of the Department of Psychology of Goldsmiths College, University of London (27th of March 2018). A field experiment in the London Underground was designed to measure the effects of music familiarity (familiar vs. unfamiliar) and performer's body movements (expressive vs. restricted). The dependent variables were the amount of money donated and the number of donors. Each session lasted approximately an hour and was comprised of four blocks: (i) familiar music with body movements, (ii) familiar music without body movements, (iii) unfamiliar music with body movements, and (iv) unfamiliar music without body movements. The order of the four blocks was fully counterbalanced across sessions using a Latin Square Design (see Berman & Fryer, 2014, for a review), resulting in a total of 24 possible orders. Figure 1 gives a graphical description of the experimental design.

[insert Figure 1.]

Experimental setup

Figure 2 shows a picture of the busker performing in one of the experimental sessions in the London Underground. The field experiment was always performed in the same location, namely, busking pitch #3 in London's Waterloo underground station. Waterloo is the busiest underground station in London, servicing 100.3 million

passengers per year (https://tfl.gov.uk/corporate/about-tfl/what-we-do/london-underground/facts-and-figures). This location was chosen primarily because the busker had previously performed there many times and as it was a relatively easy pitch to book compared to other locations. This ensured we could book the same pitch for all 24 experimental sessions. Moreover, a decision was made to conduct the field experiment in the Underground, instead of other outdoor locations, in order to be able to control for potential extraneous variables such as weather. The busker was a professional singer who has been licensed to busk in London Underground by Transport for London since 2017, when the first busking licenses were issued.

[insert Figure 2.]

To set up the session, the iPod was plugged into the auxiliary input of a Roland Cube Street battery powered amplifier, along with a Shure SM58 microphone, which was turned off to avoid sending any noise or feedback through the amp during the mime. The volume of audio output was controlled from the iPod, and the level was kept constant across all sessions. A standard metal music stand was erected, and an Akaso EK5000 video camera set to 1080p/30fps mounted on a Rhodesy Octopus-style tripod was wrapped around the pole. The busker's money collection bag, sized approximately 30cm x 60cm x 20cm, was positioned next to the music stand. The camera was aimed down at the money bag. This camera was used to record the amount of money donated as well as the number of people donating (see supplementary materials for the video footage of one of the sessions).

To determine the amount of money donated more efficiently after each block, four layers of scarves were arranged in the busker's collection bag. Each block condition

was assigned a different coloured scarf – green for familiar/expressive, blue for familiar/restricted, purple for unfamiliar/expressive, and magenta for unfamiliar/restricted. The scarf colour assigned to the last block condition of the session was placed on the bottom of the bag, followed by the penultimate block condition, until the colour ascribed to the first block condition which was placed on top. At the end of each block, the money donated by onlookers during that block was quickly scooped up in the scarf, tied up and set aside, leaving the bag empty and ready for donations to be given in the next block.

Prior to the experimental session, a pilot of the experimental setup was conducted. Two researchers (HT and DO) were present during the pilot, listening to and watching the performance, and a third researcher (MAT) reviewed the pilot session from video recordings. It was concluded that the acoustics were fitting, lip-syncing was unnoticeable, and the passersby were engaging with the performer in a typical fashion.

Music Stimuli

Stimuli Selection pre-study

In order to select music stimuli that differed only in their familiarity and were as similar as possible in other features (e.g., style, instrumentation, production), we conducted an online study using Qualtrics software (Qualtrics, Provo, UT). A total of 40 songs were chosen from 10 artists, whereby the four songs from each artist had been released in the same album. The criteria for selection were female artists (or female-fronted bands) who had had a Top 10 hit on the UK singles charts. The hit song had to be on the same album as at least three other songs that were not released as singles in the UK and had, therefore, not achieved as much popularity as the hit. Accordingly, these three songs, although similar in relevant music properties, including singer, year

of release, style, instrumentation, and production, were unlikely to be as familiar to the general public.

Table 1 shows the ten hit songs deemed as highly familiar and the three matched songs from the same artist and album that had not achieved as much popularity. Each song was trimmed to a 30 second excerpt, as close to the chorus or the most repeating (or familiar) segment of the track as possible, using the music creation software GarageBand, version 10.2.

[insert Table 1.]

A sample of 53 participants took part in the online study. Participation was on a voluntary basis and unpaid. Participants listened to the 10 hit songs from the different artists and rated how familiar each song was to them, on a scale from 1 (not at all) to 6 (very much). The order of presentation of the 10 hit songs was randomized for each participant. Along with the presentation of each hit, participants were presented with the three matched tracks from the same artist released in the same album, also in random order. They were asked to evaluate how familiar each of the three tracks was to them, using the same 6-point scale, as well as to evaluate their similarity to the hit, on scale from 1 (not at all) to 6 (very much). Participants were not prompted to consider precisely *how* the songs were similar (e.g., key, tempo, theme, chord progression, song structure). Rather, the question was left to the interpretation of the survey respondent.

Figure 3 shows the mean scores of familiarity of the ten hits and their respective three matched songs, as well as the mean scores of similarity of the matched songs to the hits. Note that although the hit highest in familiarity was by Cyndi Lauper, a decision was made to exclude this artist since the song was released in an album in

1983, whereas the other albums were from 2000s onwards. Thus, for the familiar music condition, we selected the following four highly popular (most familiar to respondents) hits: “Firework” by Katy Perry, “Stronger” by Kelly Clarkson, “Applause” by Lady Gaga, and “Sober” by Pink. For the unfamiliar music condition, matched songs by each of these four artists were selected based on their low familiarity ratings but high evaluations on similarity to the corresponding hit, namely, “Hummingbird Heartbeat” by Katy Perry, “Alone” by Kelly Clarkson, “Fashion!” by Lady Gaga, and “I Don’t Believe You” by Pink.

[insert Figure 3.]

Pre-recording and stimuli preparation

Instrumental versions of the four familiar and four matched-unfamiliar songs were downloaded online (www.youtube.com and www.karaoke-version.com). The busker’s voice was recorded using Logic Pro X recording software and a Rode NT1 microphone, creating audio versions of the busker singing on each of the eight instrumental recordings. The songs were loaded into iTunes. Two separate playlists were created, one with the four familiar songs and one with the four unfamiliar songs, so that each playlist could be played according to the block condition. An extra track consisting of five seconds of silence was added as the starting track into each playlist to ensure that the songs would randomize correctly without the need to start the playlist manually from a particular tune. The two playlists were then downloaded onto a 4GB iPod Nano A1236. The total playing time was 15 minutes and 25 seconds for the four songs

in the familiar condition, and 15 minutes and 24 seconds for the four songs in the unfamiliar condition.

Procedure

At the start of the session, the busker was reminded of the block order for the session. The layers of scarves of different colours (representing different blocks) were arranged accordingly. The investigator moved some distance away as to be as unobtrusive and inconspicuous as possible. The order of the songs in each block were played in random order using iTunes. During the experimental sessions, the busker lip-synced to pre-recorded recordings so that audio input in the experiment remained identical across sessions. Thus, the only variables that changed across conditions were the familiarity of the music and the expressivity of the body movements, which could be expressive (e.g., swaying, hand gestures) or restricted (the performer remained as still as possible), depending on the assigned condition of the block. At the end of each block, the investigator approached the busker to collect the donations in the scarf and ensure that the busker was aware of the next block condition. At the end of the session, the investigator opened the scarves containing the money and counted the currency within each one on camera, logging the amount earned in donations for each block condition. The money was then given to the busker. Footage from the field sessions was later uploaded and watched back in order to count the number of donors per block condition. The first experimental session was on the 21st of June 2018 and the last on the 2nd of August 2018.

Results

To test the main hypotheses regarding the effects of familiarity and body movements, a first analysis was conducted using a chi-square test. The frequency of donors was compared between the four experimental conditions. The results showed

that there were no significant differences in the number of donors across the four different conditions, X^2 (1) = .54, p = .46: familiar music with expressive movements (25.5%), familiar music with restricted movements (25.5%), unfamiliar music with expressive movements (22.3%), and unfamiliar music with restricted movements (26.6%).

A second analysis used liner mixed-effect modelling, as implemented in the R package *lme4* (Bates, Mächler, Bolker, & Walker, 2015), which is a more advanced statistical technique that takes into account the repeated measures structure of the data and can model random variability by assuming random intercepts for different relevant factors, such as the day of the experiment, time, and the order of the experimental blocks (Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000).

We ran separate analyses for the two dependent variables: the amount of money donated (donations) and the number of people who donated (donors). Based on Ekström (2012), the experimental sessions in a given day were taken as the repeated measure unit. In the two analyses, familiarity (familiar vs. unfamiliar music), body movements (expressive vs. restricted), and the interaction term were the fixed effect factors, while the day of the session was the random effect factor. Note that adding intercepts for order of the blocks, time of the day, week, and month did not improve the overall performance of the models and, therefore, they were not included in the final model. Effect coding (as opposed to the default treatment coding) as well as Type-III Wald chi-square significance test were used, as implemented in the R package *car* (Fox et al., 2011). Effects sizes were calculated using the R package *MuMIn* (Barton, 2009), which calculates the marginal (variance explained by the fixed factors) and the conditional (variance explained by both fixed and random factors) coefficient of determination for

Generalized mixed-effect models. See Appendix A for a summary table of the two linear mixed-effects models (donations and donors).

Figure 4 shows the effects of familiarity and body movements on the amount of money donated. The linear-mixed effect model revealed that familiarity, body movements, and the interaction term were all nonsignificant (all *p*-values > .05). The marginal and conditional effect sizes of the model were .033 and .107, respectively. Figure 5 depicts the effects of familiarity and body movements on the number of people donating money (donors). The linear-mixed effect model, again, indicated that none of the fixed factors (i.e., familiarity, body movements, and the interaction) were statistically significant (all *p*-values > .05). The marginal and conditional effect sizes of the model were .023 and .023, respectively.

Overall, in the familiar music condition, the average monetary value of donations was £3.58 (*SD*= 2.92) and the average number of donors was 2.96 (*SD*= 1.74), whereas in the unfamiliar music condition, the averages were £3.10 (*SD*= 2.71) and 2.81 donors (*SD*= 1.50). In the expressive body movements condition, the average monetary value of donations was £3.14 (*SD*= 2.66) and the average number of donors was 2.73 (*SD*= 1.57), whereas in the restricted body movements condition, the averages were £3.55 (*SD*= 2.98) and 3.04 donators (*SD*= 1.66), respectively.

[insert Figure 4.]

[insert Figure 5.]

Discussion

The present study aimed to investigate the extent to which performative aspects

(i.e., music familiarity and expressivity of body movements) influence behavioural responses to street music performances. The results from the field experiment did not support our previous hypotheses. Firstly, the familiarity of the music did not have a significant impact on the amount of donations and number of donors. This finding was initially surprising given the large amount of research showing the effects of familiarity on liking for music (see North & Hargreaves, 2008, for a review), music performances (Anglada-Tort & Müllensifen, 2017; Korger & Margulis, 2016), emotional engagement to music (Pereira et al., 2011), and willingness to pay for music (Tavani et al., 2016). This result occurred in spite of testing the stimuli in a previous online study in which we carefully selected music based on their familiarity while remaining as similar as possible in other relevant features (e.g., artist, year of release, style, instrumentation, production). Thus, our study does not support previous literature on familiarity effects and music. Alternatively, it could be argued that the magnitude of any existing effect was too small to be detected by measuring donating behaviour alone. For example, within the expressive body movements condition, there was a trend supporting the hypothesis regarding familiarity (Figures 4 and 5) - i.e., familiar music led to more donations and donors than unfamiliar music. This trend was also present in the overall results across conditions, with higher donations and donors in the familiar music condition compared to the unfamiliar blocks. Based on our data, however, there is little to suggest that street music performers should opt to use familiar music stimuli over unfamiliar to create positive affect and maximize profits.

The second hypothesis with respect to expressivity of body movements was also rejected. Expressivity did not have a significant effect on the amount of donations and number of donors. Once again, this result fails to support previous studies on the influence of visual information on music evaluation (see Platz & Kopiez, 2012, for a

review and meta-analysis study). In contrast, this finding could suggest that London commuters, in general, do not pay much attention to street music performances. A similar conclusion can be drawn from the performance of one of the world's greatest violin soloists, Joshua Bell, in the Washington Metro system, who performed classical music during 43 minutes with a Stradivarius valued at 3.5 million dollars (Service, 2007). Out of 1,097 people that passed him by, only 27 donated any money and seven stopped to listen for more than a minute, earning a total of US $32 (Service, 2007). In addition, in the context of busking and, in particular, busking in the underground, visuals might play less of a role. Indeed, the time that London passengers were exposed visually to the busker's performance in our experimental setup was limited compared to the acoustics. The busking pitch was near the bottom of an escalator, in a relatively hidden corner (Figure 2). Accordingly, passersby had sight of the busker potentially as little as 5 seconds and no more than 30 seconds. By contrast, in concert environments, where listeners are exposed to visual cues as much as auditory cues, visual information has been shown to be a prominent factor influencing the appreciation of music performances (see Platz & Kopiez, 2012, for a review). Thus, it is important to make a distinction between music street performances that happen in commuting spaces, such as the London Underground, and performances in open and more static spaces, such as city squares or parks. Visual information (e.g., the busker's body movements) is likely less influential in the former than in the latter.

Thus, the paradigm used in this study cannot be strictly compared with those in the previous literature (e.g., Castellano et al., 2008; Chapados and Levitin, 2008; Timmers et al., 2006; Vines et al. 2006; Vines et al., 2003). The current study focused on music performance within a public space and measured commuters' behavioural responses using implicit methods (i.e., counting the number of donors and amount of

donations). Previous literature on performers' body movements have rather examined classical music performances, using experimental tasks that ensure, at least to some extent, that participants pay full attention to the visual and auditory stimuli. Those previous studies have also tended to use explicit methods to measure participants' reactions (e.g., Likert scales or specific questions posted by the researchers). In order to provide a stricter comparison of this study with those conducted in lab settings, future research could include an additional experiment where videos of the busking performance are presented to informed participants. In addition, collecting qualitative data from commuters (e.g., stopping them after they have donated) would also be highly insightful.

Three additional factors could explain the observed findings. First, there are important individual differences between the amount and type of movement that performers use to express their emotional intentions (Dahl & Friberg, 2004, 2007; Vines et al., 2006; Wanderley, 2002). Second, not all performers use expressive movements in a distinct way that can be always interpreted by observers (Dahl & Friberg, 2004, 2007). Third, Wanderley (2002) reported that some of the clarinet performers under study moved while playing even when they were told not to move at all. Therefore, future studies would benefit from videoing the buskers so that an independent sample could assess the degree of expressivity in the performer's body movements. It would also be interesting to take several buskers into account in the same study to examine potential individual differences between busker types.

Our findings bring to light further interesting limitations related to the lack of control over extraneous factors. The experimental design used in the present study controlled for a number of key factors, including the music stimuli, order of the conditions within each session, day of the week and month, time, and location. In

addition, by conducting all experimental conditions in each session as well as fully counterbalancing their order across all experimental sessions using a Latin Square Design (see Berman & Fryer, 2014, for a review), our design was also robust to other potential extraneous variables (e.g., number of passersby in a given day). Nevertheless, there were some variables that did not remain under control and which could be potential sources of random noise. As discussed above, differences between the busker's body movements across experimental conditions and sessions were not systematically manipulated. Moreover, we did not control for the visual appearance of the busker, including their clothing, and evidence suggests that audiences might respond differently toward a performer based on physical appearance and appropriateness of dress (Griffiths, 2009; North & Hargreaves, 1997). Finally, another factor that may have had a potential influence related to the *bandwagon effect* (Leibenstein, 1950) – i.e., a psychological phenomenon in which people act in a certain way (e.g., donating money) primarily because other people are doing it. Here, the amount of money placed in the collection bag as well as the people donating in a specific point in time, might have signalled to the observer that others deemed the performer to be good, and vice versa (Kushner & Brooks, 2000).

Regarding our initial question - *are findings from lab studies reproducible outside the lab and under real-world conditions?* - the results reported in this study are incongruent with studies conducted in lab and artificial environments, looking at the effects of familiarity (see North & Hargreaves, 2008, for a review) and, at least partly, performer's body movements (see Platz & Kopiez, 2012, for a review). These discrepancies might be due to differences in the ecologic validity between laboratory and field studies. Laboratory experiments normally suffer from low ecological validity (i.e., the extent to which an experiment approximates the real-world situation under

study) and low external validity (i.e., the degree to which the results of the study can be generalizable beyond the research setting) (Carpenter, Harrison, & List, 2005; Reis & Judd, 2000). For instance, in the lab, participants are always aware of their participation in a scientific study and their only goal is to listen carefully to the music while evaluating it in a highly controlled and quiet environment. In contrast, the field experiment reported here offered high ecological validity: The 24 experimental sessions were conducted in a natural busking environment, under real-world conditions, and participants did not know they were part of a scientific study. When measuring economic behaviour, issues related to poor ecological validity and generalizability are taken particularly seriously by economists and behavioural scientists (Harrison & List, 2004; Levitt & List, 2007). As argued by Levitt and List (2007): “Perhaps the most fundamental question in experimental economics is whether findings from the lab are likely to provide reliable inferences outside of the laboratory” (p. 179). Overall, we hope to inspire both music psychologists and behavioural scientists to consider further ways to examine human behavioural responses to music and aesthetic stimuli in natural environments, once sufficient scientific grounding has been obtained based on lab-generated data.

The field of behavioural economics has increased the realism of the psychological underpinnings of economic analysis, improving substantially its explanatory power (Camerer & Loewenstein, 2011). Behavioural economics has not only transformed traditional economics, but it has also had far-reaching implications for many other fields, including psychology, political sciences, health, law, education, and marketing (see Angner, 2012; Cartwright, 2014; Dhami, 2016; Hastie & Dawes, 2010; Kahneman, 2011, for reviews). Nevertheless, despite the popularity of behavioural economics, the field has not yet been applied explicitly to the study of judgements and

decision-making processes in the context of music listening and music-related phenomena. Thus, *the behavioural economics of music* (Anglada-Tort & Müllensiefen, 2017; Anglada-Tort, Baker, & Müllensiefen, 2018; Anglada-Tort, Steffens, & Müllensiefen, 2018) aims to create a solid understanding of the role that behavioural economics and the psychology of decision making can play to study music judgements, choice behaviour, and aesthetics. In the present study, we applied field research methodology commonly used by behavioural scientists and experimental economists to investigate donating behaviour and charitable giving in the real world (e.g., Ebeling et al., 2017; Ekström, 2012; Khadjavi, 2016; Moussaoui et al., 2016; Olda & Ichihashi, 2016). We hope to show potential applications and benefits of *the behavioural economics of music* and encourage future researchers to apply paradigms and knowledge from behavioural economics and the psychology of decision making to study music and music-related phenomena.

Conflict of Interest Statement

The authors report no conflicts of interest.

References

Anglada-Tort, M., & Müllensiefen, D. (2017). The repeated recording illusion: The effects of extrinsic and individual difference factors on musical judgements. *Music Perception, 35(1),* 92-115. doi: 10.1525/mp.2017.35.1.94

Anglada-Tort, M., Baker, T., & Müllensiefen, D. (2018). False memories in music listening: exploring the misinformation effect and individual difference factors in auditory memory. *Memory,* 1-16. doi: 10.1080/09658211.2018.1545858

Anglada-Tort, M., Steffens, J., & Müllensiefen, D. (2018). Names and titles matter: The impact of linguistic fluency and the affect heuristic on aesthetic and value

judgements of music. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. doi: 10.1037/aca0000172

Angner, E. (2012). *A course in behavioral economics*. Macmillan International Higher Education.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. Retrieved from https://arxiv.org/pdf/1406.5823.pdf

Barton, K., & Barton, M. K. (2018). Package 'MuMIn'. R Package Version 1.42.1. Retrieved from https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59(4), 390–412.

Breyley, G. J. (2016). Between the cracks: Street music in Iran. *Journal of Musicological Research, 35*(2), 72-81.

Campbell, P. J. (1981). *Passing the hat: Street performers in America*. Delacorte Press.

Carpenter, J. P., Harrison, G. W., & List, J. A. (Eds.). (2005). *Field experiments in economics*. Elsevier JAI.

Cartwright, E. (2011). *Behavioral economics*. Routledge.

Castellano, G., Mortillaro, M., Camurri, A., Volpe, G., & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception: An Interdisciplinary Journal, 26*(2), 103-119. doi:10.1525/mp.2008.26.2.103

Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition, 108*(3), 639-651.

Cohen, D., & Greenwood, B. (1981). *The buskers: A history of street entertainment*. London, England: David and Charles.

Dahl, S., & Friberg, A. (2004). Expressiveness of a marimba player's body movements. *TMH-SPSR, 46(1),* 75-86.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception: An Interdisciplinary Journal*, *24*(5), 433-454.

Dhami, S. (2016). The foundations of behavioral economics analysis. Oxford.

Ebeling, F., Feldhaus, C., & Fendrich, J. (2017). A field experiment on the impact of a prior donor's social status on subsequent charitable giving. *Journal of Economic Psychology*, *61*, 124-133.

Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*(3), 530-546.

Foster, A. (2014, May 6). *Busking? I won't let it go, says Passenger*. Retrieved from https://www.standard.co.uk/showbiz/busking-i-won-t-let-it-go-says-passenger-9325716.html

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graces, S., Heiberger, R., Monette, G., Murdoch, D., Nilsson, H., Ogle, D., Ripley, B., Venable, W., Winsemius, D., Zeileis (2017). The car package. R Package Version 2.1-6. Retrieved from http://ftp.uni-bayreuth.de/math/statlib/R/CRAN/doc/packages/car.pdf

Griffiths, N. K. (2009). 'Posh music should equal posh dress': An investigation into the concert dress and physical appearance of female soloists. *Psychology of Music*, *38*(2), 159-177.

Hallam, s., Cross, I., & Thaut, M. (2016). *Oxford handbook of music psychology* (2nd ed.). Oxford, UK: Oxford University Press.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic literature*, *42*(4), 1009-1055.

Hastie, R., & Dawes, R. M. (2010). *Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making*. Thousand Oaks, CA: SAGE Publications.

Jacob, C., Guéguen, N., & Boulbry, G. (2010). Effects of songs with prosocial lyrics on tipping behavior in a restaurant. *International Journal of Hospitality Management*, *29*(4), 761-763.

Jeffreys, E., & Wang, S. (2012). Migrant Beegars and buskers: China's have-less celebrities. *Critical Asian Studies*, *44*(4), 571-596.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Khadjavi, M. (2016). Indirect reciprocity and charitable giving—evidence from a field experiment. *Management Science*, *63*(11), 3708-3717.

Kroger, C., & Margulis, E. H. (2016). “But they told me it was professional”: Extrinsic factors in the evaluation of musical performance. *Psychology of Music*, *45*(1), 49-64. doi: 10.1177/0305735616642543

Kushner, R. J., & Brooks, A. C. (2000). The one-man band by the quick lunch stand: Modeling audience response to street performance. *Journal of Cultural Economics*, *24*(1), 65-77.

Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *The quarterly journal of economics*, *64*(2), 183-207.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic perspectives*, *21*(2), 153-174.

Lemay, J. O., & Bates, L. W. (2013). Exploration of charity toward busking (street performance) as a function of religion. *Psychological Reports*, *112*(2), 578-592.

Marina, P. (2018). Buskers of New Orleans: Transgressive sociology in the urban underbelly. *Journal of Contemporary Ethnography*, *47*(3), 306-335.

McNamara, L., & Quilter, J. (2016). Street music and the law in Australia: busker perspectives on the impact of local council rules and regulations. *Journal of Musicological Research*, *35*(2), 113-127.

Moussaoui, L. S., Naef, D., Tissot, J. D., & Desrichard, O. (2016). "Save lives" arguments might not be as effective as you think: A randomized field experiment on blood donation. *Transfusion clinique et biologique*, *23*(2), 59-63.

North, A. C., & Hargreaves, D. J. (1997). The effect of physical attractiveness on responses to pop music performers and their music. *Empirical Studies of the Arts*, *15*(1), 75-89.

North, A., & Hargreaves, D. (2008). *The social and applied psychology of music*. New York, NY: Oxford University Press.

North, A. C., Tarrant, M., & Hargreaves, D. J. (2004). The effects of music on helping behavior: A field study. *Environment and Behavior*, *36*(2), 266-275.

Oda, R., & Ichihashi, R. (2016). Effects of eye images and norm cues on charitable donation: A field experiment in an izakaya. *Evolutionary Psychology*, *14*(4), 1474704916668874.

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PloS one*, *6*(11), e27241.

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: Basic concepts and examples. In *Mixed-effects models in S and S-plus (pp. 3–56)*. New York: Springer.

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception*, *30*(1), 71–83.

Reis, H. T., & Judd, C. M. (Eds.). (2000). *Handbook of research methods in social and personality psychology*. Cambridge University Press.

Ruth, N. (2017). "Heal the World": A field experiment on the effects of music with prosocial lyrics on prosocial behavior. *Psychology of Music*, *45*(2), 298-304.

Service, T. (2007, April 18). Joshua Bell: no ordinary busker. *The Guardian.* Retrieved from https://www.theguardian.com/music/tomserviceblog/2007/apr/18/joshuabellnoordinarybusker

Tavani, J. L., Caroff, X., Storme, M., & Collange, J. (2016). Familiarity and liking for music: The moderating effect of creative potential and what predict the market value. *Learning and Individual Differences*, *52*, 197-203.

Quilter, J., & McNamara, L. (2015). Long may the buskers carry on busking: Street music and the law in Melbourne and Sydney. *Melbourne University Law Review*, *39*, 539.

Rebeiro Gruhl, K. (2017). Becoming visible: Exploring the meaning of busking for a person with mental illness. *Journal of Occupational Science*, *24*(2), 193-202.

Smith, M. (1996). Traditions, stereotypes, and tactics: A history of musical buskers in Toronto. *Canadian Journal for Traditional Music*, *24*(6), 6-22.

Timmers, R., Marolt, M., Camurri, A., & Volpe, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: an explorative case study. *Psychology of Music*, *34*(4), 481-510.

Vines, B., Krumhansl, C., Wanderley, M., & Levitin, D. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80-113.

Vines, B. W., Wanderley, M. M., Krumhansl, C. L., Nuzzo, R. L., & Levitin, D. J. (2003). Performance gestures of musicians: What structural and emotional information do they convey? *Gesture-Based Communication in Human-Computer Interaction*, 468-478.

Wanderley, M. M. (2001). Quantitative analysis of non-obvious performer gestures. In *International Gesture Workshop* (pp. 241-253). Springer, Berlin, Heidelberg.

Williams, J. (2016). Busking in musical thought: value, affect, and becoming. *Journal of Musicological Research*, *35*(2), 142-155.

Wong, K. K. T. (2016). An ethnomusicological understanding of the street performance of Cantonese opera 街檔 (Jie Dang) in Hong Kong. *Journal of Musicological Research*, *35*(2), 100-112.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9* (2p2), 1-27.

Appendix A

Summary table of the linear mixed-effects models

Dependent variable	*Chi.sq*	*df*	*p-value*
Donations			
Familiarity	.20	1	.66
Body movements	2.52	1	.11
Interaction	2.23	1	.13
Donors			
Familiarity	.13	1	.72
Body movements	1.78	1	.18
Interaction	.89	1	.35

Table 1. Ten hit songs and the three matched songs used in the stimuli selection pre-study.

Artist *(Album, year)*	Hit	Match 1	Match 2	Match 3
Cyndi Lauper *(She's So Unusual, 1983)*	Girls Just Want to Have Fun	Money Changes Everything	When You Were Mine	Yeah Yeah
Ellie Goulding *(Lights, 2010)*	Starry-Eyed	Everytime You Go	Your Biggest Mistake	This Love
Jessie J *(Who You Are, 2011)*	Domino	Abracadabra	I Need This	Rainbow
Kelly Clarkson *(Stronger, 2012)*	Stronger	You Love Me	Hello	Alone
Kylie Minogue *(Fever, 2001)*	Come Into My World	More, More, More	Fever	Burning Up
Katy Perry *(Teenage Dream, 2010)*	Firework	Hummingbird Heartbeat	Peacock	Circle the Drain
Lady Gaga *(Artpop, 2013)*	Applause	Venus	MANiCURE	Fashion
Lily Allen *(It's Not Me, It's You, 2009)*	The Fear	I Could Say	Back to the Start	Never Gonna Happen
Little Boots *(Hands, 2009)*	Remedy	Mathematics	Tune Into My Heart	Click
Pink *(Funhouse, 2008)*	Sober	I Don't Believe You	This Is How It Goes Down	Bad Influence

Note. The hit songs charted at number nine or higher on the UK singles charts. The matched songs were matched in terms of artists and album.

Figure 1. Experimental design.

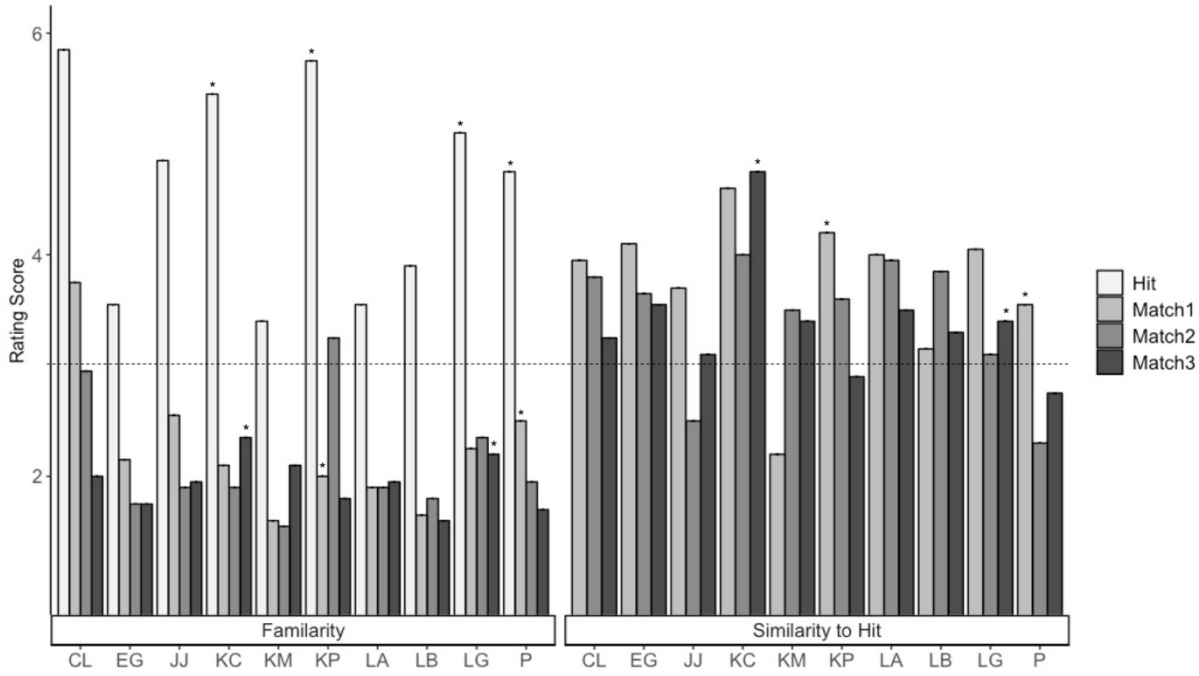
Days/ Between	Blocks/ Within 1	2	3	4
1	Familiar + Expressive	Familiar + Restricted	Unfamiliar + Expressive	Unfamiliar + Restricted
2	Familiar + Expressive	Familiar + Restricted	Unfamiliar + Restricted	Unfamiliar + Expressive
3	Familiar + Expressive	Unfamiliar + Expressive	Unfamiliar + Restricted	Familiar + Restricted
4	Familiar + Expressive	Unfamiliar + Expressive	Familiar + Restricted	Unfamiliar + Restricted
5	Familiar + Expressive	Unfamiliar + Restricted	Familiar + Restricted	Unfamiliar + Expressive
6	Familiar + Expressive	Unfamiliar + Restricted	Unfamiliar + Expressive	Familiar + Restricted
7	Familiar + Restricted	Familiar + Expressive	Unfamiliar + Expressive	Unfamiliar + Restricted
8	Familiar + Restricted	Familiar + Expressive	Unfamiliar + Restricted	Unfamiliar + Expressive
9	Familiar + Restricted	Unfamiliar + Expressive	Familiar + Expressive	Unfamiliar + Restricted
10	Familiar + Restricted	Unfamiliar + Expressive	Unfamiliar + Restricted	Familiar + Expressive
11	Familiar + Restricted	Unfamiliar + Restricted	Familiar + Expressive	Unfamiliar + Expressive
12	Familiar + Restricted	Unfamiliar + Restricted	Unfamiliar + Expressive	Familiar + Expressive
13	Unfamiliar + Expressive	Familiar + Expressive	Familiar + Restricted	Unfamiliar + Restricted
14	Unfamiliar + Expressive	Familiar + Expressive	Unfamiliar + Restricted	Familiar + Restricted
15	Unfamiliar + Expressive	Familiar + Restricted	Familiar + Expressive	Unfamiliar + Restricted
16	Unfamiliar + Expressive	Familiar + Restricted	Unfamiliar + Restricted	Familiar + Expressive
17	Unfamiliar + Expressive	Unfamiliar + Restricted	Familiar + Expressive	Familiar + Restricted
18	Unfamiliar + Expressive	Unfamiliar + Restricted	Familiar + Restricted	Familiar + Expressive
19	Unfamiliar + Restricted	Familiar + Expressive	Familiar + Restricted	Unfamiliar + Expressive
20	Unfamiliar + Restricted	Familiar + Expressive	Unfamiliar + Expressive	Familiar + Restricted
21	Unfamiliar + Restricted	Unfamiliar + Expressive	Familiar + Expressive	Familiar + Restricted
22	Unfamiliar + Restricted	Unfamiliar + Expressive	Familiar + Restricted	Familiar + Expressive
23	Unfamiliar + Restricted	Familiar + Restricted	Familiar + Expressive	Unfamiliar + Expressive
24	Unfamiliar + Restricted	Familiar + Restricted	Unfamiliar + Expressive	Familiar + Expressive

Note. The order of the four blocks was fully counterbalanced across sessions using a Latin Square Design, resulting in a total of 24 possible orders.

Figure 2. Picture of the busker performing at pitch #3 in London's Waterloo underground station.

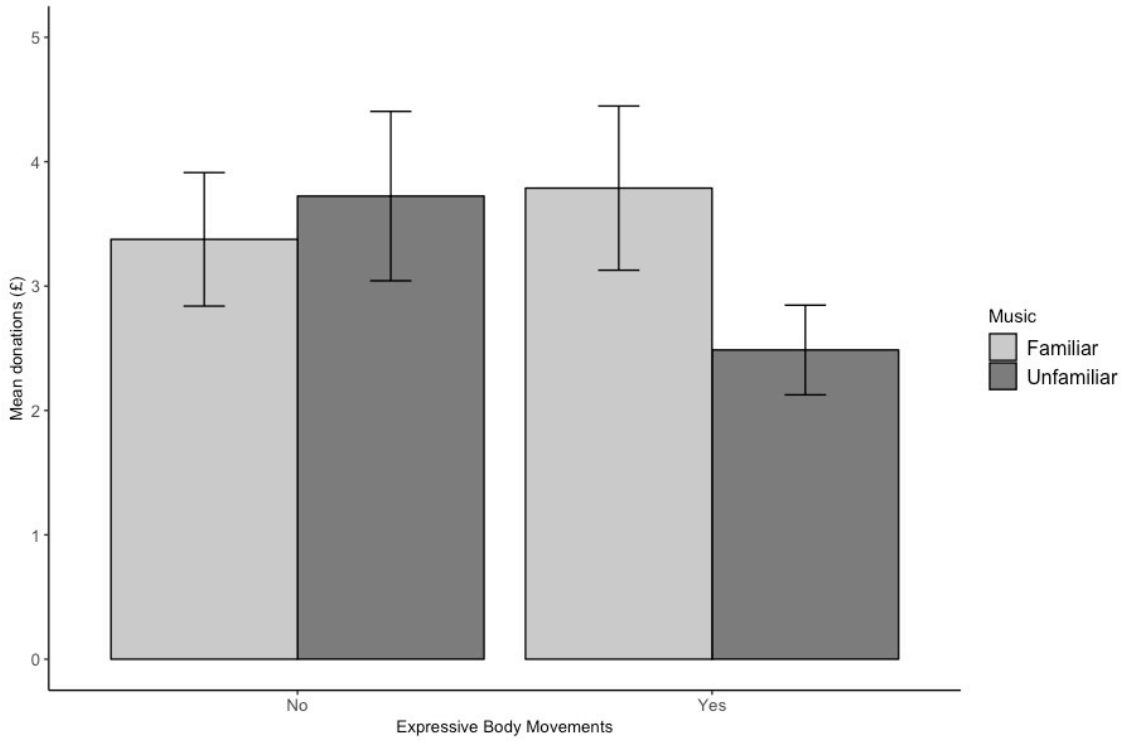
Note. This was the experimental setup used in the 24 sessions. The camera, which records the amount of money donated and number of people donating, is mounted on the stand. In the bottom right area of the pitch, is the locations of the money bag where donations are solicited.

Figure 3. Mean familiarity of the ten hits and three respective matched songs (by artist and album) and mean similarity of each of the matched songs to the respective hits.



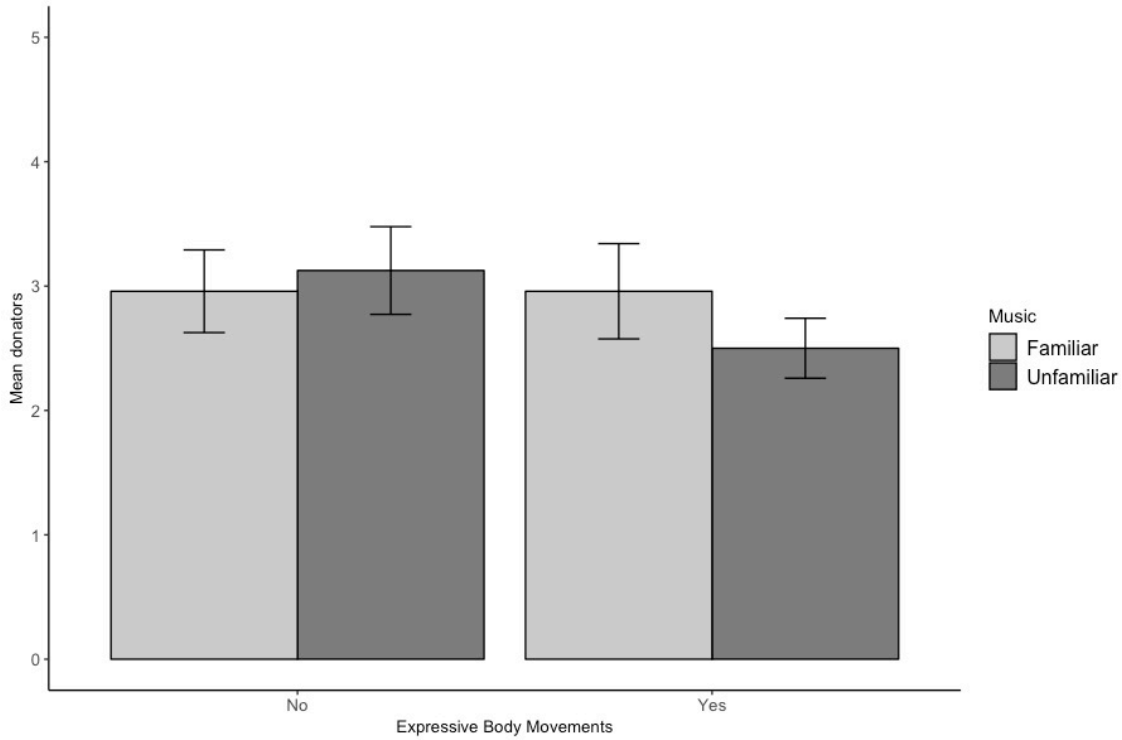
Note. CL= Cyndi Lauper; EG= Ellie Goulding; JJ= Jessie J; KC= Kelly Clarkson; KM= Kylie Minogue; KP= Katy Perry; LA= Lily Allen; LG= Lady Gaga; LB= Little Boots; P= Pink. The asterisks (*) indicate the four hits selected based on high familiarity scores and the four corresponding matched songs, selected based on low familiarity scores and high ratings of similarity to hit. The dashed line indicates the middle of the 6-point scale.

Figure 4. Effects of familiarity and body movements on the amount of money donated



Note. Error bars represent the standard error.

Figure 5. Effects of familiarity and body movements on the number of donors



Note. Error bars represent the standard error.