



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale

Corso di laurea in Scienze Psicologiche Cognitive e Psicobiologiche

Elaborato finale

The Music of Power: a direct replication study

Relatore:

Prof. Massimo Grassi

***Laureanda:* Silvia Buratto**

***Matricola:* 1221890**

Anno Accademico 2021/2022

Index

Index	
1. Introduction	1
2. The replicability crisis	2
2.1 An historical overview	2
2.2 The replicability crisis in social psychology	3
2.3 To improve replicability	5
3. The case of Power Posing	7
4. The Music of Power	10
5. A direct replication study	11
5.1 Pretests	11
5.1.1 Results and discussion.....	12
5.2 Replication of Experiment 3	13
5.2.1 Participants	13
5.2.2 Procedure	13
5.2.3 Instruments	14
5.2.4 Measures	14
5.3 Results	15
5.4 Discussion	15
5.4.1 Changes to the original paradigm.....	15
5.4.2 Criticalities of my research project.....	17
6. Conclusion	19
7. References.....	22
Acknowledgements.....	
Appendix.....	
Pre-registration module on Aspredicted.org.....	
Songs used for the pretests.....	
a. Complete list of songs in Spotify’s playlist Top Hits Italia (02/02/2022) ...	

b. List of the 31 songs pretested
Results of the pretests
a. SONGS – pretest
b. LYRICS – pretest
Die Rolling Task

1. Introduction

In this thesis I will first provide an overview of the replicability crisis in psychology, with its development in recent years, and proceed to examine it in relation to social psychology. I will then bring forward a prominent example of the debate around replicability, which occurred to power posing (Carney, Cuddy, & Yap, 2010). This effect, that can be located within social psychology and in particular among the works on the psychology of power, is a thematically near phenomenon to the paper I chose to replicate. Therefore, I will then draw a connection to “The Music of Power: Perceptual and Behavioral Consequences of Powerful Music” (Hsu, Huang, Nordgren, Rucker & Galinsky, 2015) and proceed to give a summary of it. After a methodological description of the pretests and the replicated experiment, I will report the results and conclude the work by discussing the role of the changes in the procedure and the logic behind them. I will then end with a comparison between my results and the original study, exploring the possible critical factors that may have determined any different outcome.

2. The replicability crisis

2.1 An historical overview

In the last years psychology, alongside many other sciences, like economy, medicine, and philosophy, has faced what is known as the replicability crisis. Publications, which formed the theoretical bases for entire generations of psychologists, have undergone replication efforts, often with no success in confirming the original results.

Although this phenomenon has already been in the spotlight of the discipline in the past century, it has reached the state of a proper crisis only in the last decade. To give a brief temporal reference, this debate took its origins from three main events, as exposed by Pashler and Wagenmakers (2012) and Wiggins and Christopherson (2019). They identified the publication of an article titled “Feeling the future” in the “Journal of Social and Personality Psychology” (Bem, 2011) as the first critical step. When this article came out, it apparently provided evidence for the human ability of precognition or the capacity to foretell future events. The community took this as an alarm, highlighting that, if a prominent journal is willing to publish findings on presumably unobtainable phenomena, there is a problem with research practices as a whole and with the publication standards.

In the same year, more doubts grew after the Diederik Stapel fraud case (see Stroebe, Postmes, & Spears, 2012 in Pashler & Wagenmakers, 2012), regarding one famous social psychologist, with a long list of works published in high profile journals, who was found guilty of fabricating data and therefore forcing the results of many of the researches in question. This event shed another light on the appropriateness of the verification processes to which experimental psychology was subject.

Lastly, a team of authors formed by Simmons, Nelson, and Simonsohn (2011) published an article about how methodological distortions were likely responsible for the large amount of significant results in psychology. They were among the first to demonstrate with reliable evidence that employing such techniques, also labeled as Questionable Research Practices (QRPs), allowed

many studies to force positive results (Pashler & Wagenmakers, 2012), hence to confirm the initial hypotheses in the absence of a real effect.

The term “*replicability crisis*” does not only refer to the events of 2012, but it takes its origin from them. In fact, these publications and qualitatively problematic practices started a more discipline-wide and systematic series of reviews. The debate unfolded in such a way that led to replication attempts around publications, which represented the foundations of the whole subject. Significant results have not been successfully replicated in the majority of the examined cases. To name a couple of examples, ego-depletion, which refers to a model of self-control as a finite quantity that can be consumed (Baumeister, Bratslavsky, Muraven, & Tice, 1998) and unconscious priming effects (Bargh, Chen, & Burrows, 1996) were both subjected to replications, which failed to confirm the previously observed effects. (Świątkowski & Dompnier, 2017)

To summarize, when we talk about the replicability crisis we are not only talking about the events just discussed, but also about the large-scale failure to replicate key findings of the discipline, which led to the realization of some intrinsic weaknesses of the research processes employed throughout the history of psychology. Therefore, this phenomenon is often also depicted as a “Confidence Crisis” (Pashler & Wagenmakers, 2012), because one of the premises of science in general, is that the same procedure and experimental methods should yield the same results. For this reason, not having reliable literature and not being able to reproduce published articles mines the validity of the whole subject, substantially reducing the confidence we can have in it.

2.2 The replicability crisis in social psychology

As anticipated, after the publication of Bem’s article (2011), the discovery of scientific frauds and sub-optimal research habits, psychology has undergone a process of revision, inspired by the idea of quantifying the incidence of the problem and laying the foundations to solve it.

Within this framework, a crucial cumulative replication effort known as the “Replicability Project”, conducted by the Open Science Collaboration [OSC] (2012), aimed to quantify the percentage of works in psychological science, which held true after systematic replications. Particular emphasis was given to

recreating the conditions in which the studies were originally performed and in respecting the original paradigm, by collaborating with the authors themselves and accounting for interfering factors.

This process of carefully repeating an experimental procedure in order to “verify a piece of knowledge” takes the name of direct replication (Schmidt, 2016, p. 92, 93 in OSC, 2012). For further clarity, there are two types of replications: direct ones and conceptual ones. While the former (direct) seek to follow closely the original procedure in such a way that would yield the same results, the latter (conceptual) often take only an inspiration from the original paradigm of a study, while expanding it, involving other measures or employing new methods to examine the same constructs. Experts argue that direct replications, although rarer, are the optimal test of a phenomenon’s validity (Pashler & Harris, 2012). A failed direct replication can help to correct any falsely significant result, but at the same time it is not likely to be published (Schimmack, 2020). Conceptual replications, on the other hand, are carried out with a higher frequency probably because of how convenient they are in relation to their publication-worthiness. In fact, a successful exploratory work, which yields a new significant effect, can gain more interest, than a simple replication of an already existing effect. This reasoning falls short in considering that undertaking only conceptual replications, actually increases the impact of publication bias, discouraging any scientist to pursue direct replications. Thus, the ultimate result of this mechanism pushes psychologists to question less the existing results and to avoid any direct replica, since it represents only a waste of resources.

Coming back to the Reproducibility Project, they did not fear the obstacles against direct replications and selected a body of publications and a strict protocol to follow, in order to try and uncover confirming evidence. The team decided to take into account the whole subject and at the same time also separately analyzed two sub-fields: cognitive psychology (selecting publications from the “Journal of Experimental Psychology: Learning, Memory, and Cognition”) and Social Psychology (extracting the articles to replicate from the “Journal of Personality and Social Psychology”).

While less than half (37%) of the examined studies across all the areas managed to be replicated successfully, the success rate for social psychology

was much lower. This result has a particular relevance for this thesis, since the study I attempted to replicate falls within this field. The calculations pointed out that only one quarter (25%) of the effects replicated a significant result (Open Science Collaboration, 2015). The difference observed with cognitive psychology, which scored a way higher rate of replications (50%), is likely attributable to some factors, besides the ones previously discussed, and in the following paragraph I will explore the main ones. This next section will also explain some of the decisions that I adopted for my direct replication.

2.3 To improve replicability

The final part of this theoretical introduction wants to mention some of the suggested solutions and principles which should guide psychological research works to make findings more reliable. As Schimmack (2020) explains, many of the experiments in social psychology have low statistical power and employ between-subjects designs, which make the field more prone to failures in replications. Adopting more stable within-subjects designs, as cognitive psychology often does, is a first important step to improve the reliability of the field, together with increasing the study's sample size, which would produce higher powered designs and less false-positive results (*ibidem*).

Another critical factor for fixing the low confidence we can have in social psychological studies is to follow the principles of "Open science". This term is used as an umbrella one that refers to all the practices, which aim to reform psychological research and get to a more transparent, open and reproducible science (Crüwell et al., 2018). The key solutions according to this mentality are pre-registration and registered reports, open access to data and in general more transparency in the research process.

Pre-registering a study means that, before collecting any data, the researchers write down and store in a public digital archive all the analyses that they will lead, the number of observations they settled for, as well as all the variables included. All these relevant details are therefore already defined and the article will not run the risk of being subjected to publication biases like *file-drawering* (Rosenthal, 1979) (whereby only the experiments which uncovered a significant result are included in the publication) or selective reporting (only talking about the relevant variables, excluding the other ones which did not produce

significant effects). The pre-registration of an experimental design also helps to control for the risk of being overly flexible with the sample size, i.e. deciding it only after interim data analysis to check if the results collected are statistically significant or if more answers should be collected.

Open Access can refer to materials, data and analytic codes (Crüwell et al., 2018). This idea implies that the mentioned resources are either publicly available in online platforms or readily shared upon demand to other qualified professionals. Sharing the materials employed and the data obtained (or analyses run) allows other researchers to verify the results and to reproduce the experiment as closely as possible to the original.

In order to follow the idea of a more reliable science, this thesis includes an Appendix section, where the relevant materials are reported and includes a link to the pre-registration form on the online platform AsPredicted.org, together with a critical review of the changes from the original paradigm and the main data obtained.

This introductory section aimed to give a context within psychology to the reasons behind my decision to replicate a study. In the following paragraphs, I will first explain the process behind examining reproducibility of power posing (Carney et al., 2010) and then proceed to talk about my own work.

3. The case of Power Posing

In this next section, I will analyze more in detail a prominent case of replication failure in social psychology. The literature of power has a relatively short tradition in psychology, but it grew in the last years, especially after one of the most well-known papers on the embodied consequences of power has been published. The work by Carney, Cuddy and Yap (2010) on *power posing* gained a relevant spot both for people in the field and for the whole community. In particular, the mentioned study wanted to examine the hypothesis that poses characterized by open and expansive postures (i.e. *power poses*) were capable of creating a sensation of power. To verify it, the researchers assigned the participants to one of two conditions. The former was a power-inducing condition, where they were instructed to sit back on a chair and have their feet on the desk in front of them or to stand and lean forward on the desk, keeping their arms wide open on its surface. The latter, instead, was aimed at making the participants feel less powerful, by instructing them to sit in a chair with their arms closely together on their lap or to stand with all their limbs crossed. Before and after holding the pose, the blood levels of two hormones were measured: testosterone, related to dominance, and cortisol, related to stress, thinking that the participants in low-power positions would show increased levels of cortisol as a sign of submissiveness, while expansive positions would increase those of testosterone. The results led to stating that power-inducing poses (as opposed to closed ones) can induce higher levels of testosterone and decrease those of cortisol, while the opposite is true for the low-power poses. Furthermore, on a behavioral and cognitive level, the team found a correlation of the powerful postures with higher reported levels of power and tolerance for risks.

A central role in the replicability debate for this phenomenon is occupied by a critical replication attempt, which partially failed to confirm the phenomenon (Ranehill et al., 2015). The conceptual replication employed a higher number of subjects, for more reliable results, as well as a longer duration of the pose to hold, supposing that keeping the posture for longer would strengthen the effect. Despite adopting some critical changes to the experimental paradigm, this study successfully replicated the effects of expansive postures in increasing the

subjective feeling of power, while failing to find any effect on a physiological (hormonal) level.

To reply to this failed replication, the authors who first identified this effect wrote a narrative review of 33 works from the existing literature on power posing and examined not only the works which followed theirs, but also those published prior. They stated that the differences among the employed methods, have played a decisive role in obtaining contrasting results (Carney, Cuddy & Yap, 2015). According to them, the analyzed results show additional variables that must be considered, which may alter the validity of power posing.

Most importantly, they point out two critical aspects for the failed replication by Ranehill et al. (2015). According to them, a significant change is represented by the awareness of the participants about the tested hypothesis. While Carney et al. (2010) used a cover story to deceive the real purpose of the experiment, Ranehill and her team (2015) decided not to employ any. Secondly, the research also significantly changed the length of the posture manipulation, therefore Carney et al. (2015) argue that: "Although it may make intuitive sense that longer time in the posture would increase effects, holding some postures for too long may cause discomfort, become awkward, or habituate a body to the effects of the posture" and suggest that the effect of this variable should be directly tested in the future.

Further articles have attempted replications and a relevant pre-registered study, found opposite results, despite following the original experimental procedure more closely than previous works. The expansive posture not only did not correlate with higher risk taking, but also negatively correlated with feelings of power (Garrison, Tang, & Schmeichel, 2016).

Another prominent discussion around Power Posing and its validity, considers the review just mentioned and highlights some methodological errors (Simmons & Simonsohn, 2017). These authors used a technique, known as P-curve analysis, on the supporting studies in Carney et al. (2015)'s review. This new method was developed in response to the crisis itself and it analyzes, not without some weaknesses, the "evidential value" of sets of studies by employing

“the distribution of statistically significant p -values”, as defined by Renkewitz and Heene (2019).

They found out that the 33 studies, providing support to power posing, are subject to selective reporting effects, which means that only significant results tend to be published, as explained by Pashler and Harris (2012), and this may cause distortions. One form of selective reporting is represented by *p-hacking* (Simonsohn, Nelson, & Simmons, 2014), whereby the authors decide to conduct different or new analyses on their data until they spot a significant result. Another expression of this bias is known as *file-drawering* (Rosenthal, 1979) and consists in only including in the publications the studies, which achieved a statistically significant result.

Replies to this meta-analysis have considered a broader list of articles, aiming to use the same technique on “the best available evidence to test clearly specified a priori research hypotheses regarding well-defined effects” (Cuddy, Schultz & Fosse, 2018). These analyses came to different conclusions from those of Simonsohn et al. (2014), namely highlighting that subjective feelings of power cannot be excluded from this kind of work and that there is in fact a good evidential weight for the studies included in the p -curve technique.

The process of replication applied to this specific effect aims to be a guide for the reader. In fact, power posing is perhaps the most representative phenomenon within the field of the psychology of power which has faced the reproducibility debate, but surely not the only one. Keeping this in mind, the next section will expose the experimental procedure of the study by Hsu et al. (2015) that was the object of the replication work explained in this thesis.

4. The Music of Power

This thesis project takes its origins from a 2015 publication by Hsu and colleagues titled “*The Music of Power: Perceptual and Behavioral Consequences of Powerful Music*” published in “Social Psychological and Personality Science”. This study consists in an exploration of the hypothesis that music can have a measurable effect on the dimension of power. Through 5 experiments, the research team examined the influence of musical pieces on the implicit activation of the construct of power (Experiment 1), then they explored how a music-based sense of power related to three main consequences of power. The first cognitive consequence observed was abstract thinking (Experiment 2), which refers to the tendency to consider the whole before its parts (Smith & Trope, 2006). The following experiment focused on another cognitive consequence of power, known as illusory control (Experiment 3) or the perception of increased control over future events, which are outside of the person’s reach (Fast, Gruenfeld, Sivanathan, & Galinsky, 2009). The last power-related behavioral outcome examined was the tendency to move first (Magee, Galinsky, & Gruenfeld, 2007) (Experiment 4). The remaining two experiments tried to trace a causal link between the role of a specific feature of music, the level of bass, and the strength of the music-induced sense of power (Experiments 5a and 5b).

The first four experiments selected the stimuli able to induce power by pretesting 30-seconds song excerpts, to be high or low in power. 75 undergraduates listened to 31 excerpts and rated on a 7-point scale how *powerful*, *dominant* and *determined* each made them feel. The same procedure was used to ask 36 students to evaluate exclusively the lyrics of the same musical pieces.

This way, the studies made an effort to rule out any other potential explanation for the observed effect of music, by controlling for the effect of lyrics in the pretest phase, as described above, and removing them completely in the case of experiments 5a and 5b, where “generic instrumental music pieces” (Hsu et al., 2015) were used. Finally, researchers included a measure of positive emotions, to exclude that the effects observed are due to the association of music with the experience of such feelings.

5. A direct replication study

I decided to center my work on Experiment 3, about the relationship between music and illusory control. I chose this experiment because the measure employed does not require an elaborate translation and presents a more complete manipulation because of the association with a filler task. The pretests that were needed to define the experimental conditions were also part of my thesis project. The whole replication procedure was pre-registered on AsPredicted.org. [Appendix]

5.1 Pretests

I selected the stimuli for the power manipulation through a pretest, to identify the three most and least power-inducing songs from a list of 31 pieces. A control pretest for a potential interfering effect of the lyrics was also carried out.

I used the first 50 songs from the Spotify chart “Top Hits Italia” as on February 2nd, 2022 and the final list of the 31 songs to pretest was created by taking them in the same order as they appeared on the chart, excluding the songs, whose main author already appeared in a song in the list. [Appendix]

I proceeded to define a new selection criteria for the 30-seconds excerpts, since a bibliographical research of experiments with a comparable selection criteria for the excerpts did not produce any relevant results. I identified the starting point of the first refrain and then took the 15 seconds preceding and following that time tag. The lyrics portion corresponding to those 30 seconds, to use in the second pretest, was then selected for each song.

I built two Google Forms modules, the former consisting of 7-point (from 1=*Not at all* to 7=*Very much so*) Likert scales to collect the evaluations of the 30-seconds excerpts the participant would listen to, the latter identical to the first, except it contained a picture of the lyrics to evaluate. To rule out any effects of the order of presentation, the order of the songs was randomized. In both the pretest modules, each song evaluation was followed by a question measuring the familiarity with the song, in the form of a closed question about whether the participant knew the song or did not (*Yes, I am not sure, No*).

5.1.1 Results and discussion

The responses of 31 people (14 males, $M_{age} = 24.3$, $SD = 8.57$), who listened to the 31 excerpts physically in the lab revealed that the three highest rated songs and lowest rated ones, were those reported in Table 1.

Table 1. Highest and lowest rated music pieces (used, not corrected)

High-Power music pieces	Low-Power music pieces
1. abcdefu – GAYLE ($M = 5.67$, $SD = 1.99$)	1. Angelina Jolie – Bresh, SHUNE ($M = 2.32$, $SD = 1.34$)
2. LA PROVINCE #1 – Rhove ($M = 5.45$, $SD = 1.41$)	2. ∞ LOVE – Marracash, Guè ($M = 2.57$, $SD = 1.34$)
3. La coda del diavolo – Rkomi, Elodie ($M = 5.07$, $SD = 1.20$)	3. STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber ($M = 2.58$, $SD = 1.35$)

A further check of the analyses, carried out after the start of the administration of the experiment itself, revealed that an error occurred when calculating the means for each song, therefore two other musical pieces would have been part of the high and low power conditions instead of “*abcdefu – GAYLE*” and “*Angelina Jolie – Bresh, SHUNE*”. Table 2 reports the songs selected with the corrected calculations, which should have been used. Despite this error, a t-test confirmed that the songs used for the experiment still have a significant difference, so the error should not have influenced the manipulation.

A complete data set of the means and standard deviations used for the analyses can be found in the Appendix.

Table 2. Highest and lowest rated music pieces (corrected calculations)

High-Power music pieces	Low-Power music pieces
1. LA PROVINCE #1 – Rhove ($M = 5.45$, $SD = 1.41$)	1. ∞ LOVE – Marracash, Guè ($M = 2.57$, $SD = 1.34$)
2. La coda del diavolo – Rkomi, Elodie ($M = 5.07$, $SD = 1.20$)	2. STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber ($M = 2.58$, $SD = 1.35$)

3. LONDRA – Rosa Chemical, Rkomi $M= 4.94, SD= 1.54$)	3. SAPORE – with Tedua – Fedez, Tedua ($M= 2.61, SD= 1.15$)
---	--

To exclude that the difference between the high- and low-power pieces observed was not only due to a semantic priming effect, I compared the three songs in each of the two conditions, from the pretest with the excerpt hearing, with the same songs as they were evaluated only on the basis of the lyrics. A paired sample t-test showed a significant effect, with a higher mean for high-power songs as opposed to the low-power ones $t(33)= 2.34, p=.025$. For this reason the effect of inducing power for these pieces can be attributed to the activation determined by the words of each song.

To verify if the familiarity had any effect on the power evaluation of the songs and the lyrics, I considered whether there was a difference between the people who knew the songs and those who did not (All the respondents who chose “*I am not sure*” were merged with those who answered “*No*”). Multiple independent t-tests on the answers of 34 people (12 males, $M_{age}= 21.06, SD= 1.35$) revealed an effect of familiarity in four cases. The evaluations of the songs, which the participants listened to were different when they knew the song for “*∞ LOVE (feat. Guè) – Marracash, Guè*” $t(28)= 2.98, p= .006$ and “*Pastello Bianco – Pinguini Tattici Nucleari*” $t(28)= 2.60, p= .015$.

For what concerns the ratings received by the lyrics, “*Cold Heart – PNAU Remix – Elton John, Dua Lipa, PNAU*” $t(32)= 2.53, p= .016$ and “*Pepas – Farruko*” $t(32)= 2.12, p= .042$ were influenced by familiarity.

5.2 Replication of Experiment 3

5.2.1 Participants

30 people (10 males, $M_{age}= 22, SD= 1.53$) were randomly assigned to a high-power or low-power music condition. The order of presentation of the songs within each condition was randomized and counterbalanced.

5.2.2 Procedure

Participants were asked to sit down and received some general indications about the coming task to execute. The experimenter instructed them to fill in the

personal information section and an informed consent, then made sure that they wore the headphones correctly and started the execution of the music and the experiment. To assure that the music was played long enough to exert an effect, participants completed two filler tasks, the die-rolling task and the positive mood check while the musical pieces from the assigned condition were playing in their headphones.

5.2.3 Instruments

The experiment took place in a silent room lab in the building of the General Psychology Department of the University of Padova. The stimuli for the test were presented on a NEC monitor and an Acer computer. The audio was reproduced through a Focusrite sound card from the experimenter's laptop, Acer Swift 5 and the participant listened to the music pieces through a headphone set.

5.2.4 Measures

Illusory Control

Illusory control was measured using the die-rolling paradigm by Fast et al. (2009). Participants were asked to imagine that they could win a small sum of money by correctly predicting the outcome of a six-sided die roll. Then they were asked to express whether they would want to roll the die themselves or let the experimenter do it for them. The task was translated into Italian, the main language of the experiment. [Appendix]

The idea behind the task is that choosing to roll the die personally expresses an illusion of increased control (*illusory control*) on the results of the die toss, therefore, participants in the high-power music condition are expected to choose this option more frequently, than people in the low-power condition. The execution of the die-rolling paradigm was preceded by a filler task, where the participants had to provide five statements to describe two neutral objects (here: a chair and a lamp).

Positive emotions

Participants indicated how happy, excited and enthusiastic they felt during the execution of the task (Fast et al., 2009).

5.3 Results

There was no significant difference between the responses of participants in the high power condition (80%) and the low power condition (86.7%) $X^2(1, N=30) = .24, p = .62, \phi = .09$.

I first verified the correspondence of the three emotion ratings, to see if I could proceed to merge them. Since the correspondence levels of the evaluations on each of the three dimensions were high ($\alpha \approx .90$) I created a new variable and considered the positive emotions as a whole. I ran another t-test to exclude an effect of the conditions (high vs. low power) on positive emotions, which showed a non significant result $t(28) = .49, p = .63$.

Despite the non-significance of the results of the main experiment, I controlled for the influence of the assigned condition on positive emotions and this result also came out as non significant $B = 1.61, SE = .49, Wald = 10.79, p = .76, R^2 = .018$.

5.4 Discussion

My Bachelor thesis project aimed to replicate a published paper, to examine the relevant challenges in this process and try to verify the validity of the chosen article. I followed the original paradigm by Hsu and collaborators (2015) as closely as possible, but I also had to modify some aspects and adapt some others. The pretests procedure gave me three musical pieces for each experimental condition. The replication of the third experiment in the paper yielded some interesting results. The answers to the die rolling task in the two music conditions did not differ significantly, thus I did not find a confirmation to the existence of a music induced sense of power or to its influence on the construct of Illusory control. My analyses, on the contrary, agree with those of the original work "The Music of Power" in excluding that any effect might be attributed to positive emotions generated by the music heard.

5.4.1 Changes to the original paradigm

Since some of the key pieces of information for my replication were not available, I had to take some arbitrary decisions to be able to conduct my direct replication. The adjustments to the experimental procedure that I adopted were

led by an evaluation of which modifications were expected to have an impact and which did not.

As anticipated, I knew that the original work selected the pieces to pretest from a range of genres (i.e. hip-hop, heavy metal, sports music, reggae and punk). The unavailability of a list and lack of an answer from the authors obliged me to pick a different source for the songs for the experimental manipulation. I decided to use a chart with the 50 songs which received the highest number of streams on Spotify in Italy, in the week in which I started the creation of the materials. This change, despite allowing me to also examine the validity of the results with musical pieces relevant for the current replication environment, certainly represents a critical difference in my research project.

I assumed that in this replication attempt my choice would cut off some genres, since mainstream charts often tend to include pieces that, on one hand, stem from different genres, but which, on the other hand, actually have similar features and therefore meet the taste of the general public. The involved music genres represent a big difference between my replication and the original experiment. However, if the selection procedure is reliable enough, the rating received by the songs, should be able to identify power-inducing and non power-inducing songs, independently from the genre. My data and analyses seem to confirm this significant difference between the two conditions.

Extracting the songs from a chart including mainstream music is also interesting in terms of ecological validity of the results. In fact, it is more probable that people will hear music from such charts everyday in public places, rather than specific outdated songs, in this case perhaps more appropriate abroad than in the setting in which I replicated the study.

I introduced a new variable in the paradigm, to assess the familiarity with the songs. Controlling for it seemed a reasonable step, since charts are built off the general preference. This new variable did have an impact on 4 of the ratings, which in fact differ controlling whether the participants already knew the song or not. I already discussed this in detail in the dedicated section, but finding a difference in the power evaluations when the participant knew the songs or the lyrics could mean that this might have played a relevant role also in the inspiring

paper for this thesis. I was surprised that this factor was not accounted for in the original study, specifically because the songs which were selected and listed in the Appendix are mostly well known songs. The authors themselves open their paper by talking about the evidence relating previous psychological experiences with the exposure to music (De Houwer, Thomas, & Baeyens, 2001 in Hsu et al., 2015) and they mention a relationship between musical pieces and power, which is actually mediated by a conditioning mechanism and not by the music itself. Thus, ignoring this factor seems a questionable decision. The choice of using a popular sport-events song in the high power manipulation, for instance, could have reminded a sports fan of all the victories of the team they root for and, only through this, have evoked a power sensation.

Another relevant factor about my different music choice is the lyrics' language. While, from what can be understood in the original report by Hsu et al. (2015), they only employed instrumental pieces or songs with texts in the English language, in my case the songs were in English, Italian, Spanish and even in the Neapolitan dialect. I did not expect this factor to have any impact, since the usual repertoire of songs in the Italian mainstream music channels presents this mixture. For this reason, employing songs in multiple languages can represent a further detail increasing the ecological validity of the manipulation for the contemporary Italian context of my research project.

5.4.2 Criticalities of my research project

This section will point out two further changes I made to my experimental paradigm. The first stemmed from a practical need, whereas the other was due to a later discovered mistake.

In line with the ideas of an open and replicable psychological science, I decided to preregister my study before conducting it, but I had to make a change in the number of participants. I narrowed down the sample size of the experiment from the declared $N= 75$ to a more realistic number (considering my recruitment policy which offered no incentives to encourage participation) of around 30 people, both for the song-listening pretest and the replication of experiment 3. The new sample size was chosen to still be able to obtain stable statistical results, even with a comparatively underpowered sample.

A weakness of my research project is represented by an error I committed in running the analysis on the pretests. I did not notice, until a later reexamination, that I had used a different set of data to extract the means and standard deviations of some songs, which led me to falsely selecting 1 out of 3 songs in each of the two conditions. Although, as previously declared in the dedicated section, the analyses confirmed that the manipulation kept its strength and validity in spite of the mistake, it is an important and unpredicted change from the initial paradigm. The significant difference between the two conditions can be a sufficient proof of the reliability of my replication, whereby the manipulation is still expected to have an effect, but it still did not yield confirming results.

6. Conclusion

My final observations for this thesis are centered around the actual results of my work and what they allow me to conclude about the replicability crisis. I did not manage to confirm the findings of Hsu et al. (2015). It is still questionable to what this failure can be attributed.

The first possible cause may be the problems in the diffusion of the details behind “The Music of Power”. The imprecise description of the procedure and the other necessary aspects to conduct a replication may have pushed me to interpreting the article erroneously, changing the paradigm too much for the original effect to hold true and to consider mine a direct replication.

Secondly, the methodological problems and arbitrary decisions in my replication themselves have introduced new variables which probably altered the efficacy of the manipulation. Employing modern music in different languages may have modified the design to an extent where the effect could not be reproduced. The reduced sample size, smaller than the original one, could make this replication failure fall in the type II error and be a false negative, because the study could be unable to detect an existing effect (Schimmack, 2020).

Third, it is possible that the observed phenomenon of a music-induced sense of power is particularly influenced by *context sensitivity*, as for many other effects in social psychology. This label is used for the dependence of an effect on the environment in which phenomena are observed (Van Bavel, Mende, Siedlecki, Brady & Reiner, 2016 in Schimmack, 2020). I think this is an interesting point, since it can both represent a solution and a further weakness for the field.

Adopting a scientific practice, which is aware of context sensitivity, implies that the conditions in which the effect can be expected to occur should be accurately described. This does not mean that “Every hypothesis could be regarded as being limited by a specific set of “perspectives”, outside of which the hypothesis would no longer hold.” (Świątkowski and Dompnier, 2017) because this factor actually does not explain the low rate of replication and rather mediates as an important predictor without being fully responsible for the phenomenon (Inbar, 2016 in Schimmack, 2020). Therefore as Świątkowski and Dompnier (2017) propose, embracing Perspectivism (the philosophical-epistemological approach

deriving from context sensitivity, theorized by McGuire in 1983) in social psychology could both increase its validity and worsen the confidence in this field. It could be regarded as a positive thing, because recognizing the mediation of all the factors, that change from one social setting to another, would help the scientific community to know whether to expect a certain effect in spite of them. For example, some of the relevant variables to define could be different cultural norms or changes comparable to the ones I had to make to Hsu et al. 's (2015) paradigm. On the other hand, embracing Perspectivism may be symptomatic of a lack of generalizability of social psychological findings. This would actually accentuate the problem of the replication crisis, because the ideal context for each effect is only rarely defined and only few phenomena can be verified.

This replication attempt confronted me with a long list of weaknesses in the way the original research report is structured. I would assume that in order for a publication to be useful and in line with the idea of science itself, the description of the experimental procedure would allow every willing person to conduct a direct replication of it. When it comes to experimental psychology specifically, all the sections in a paper are expected to provide all the necessary details about the sample, measured variables, used materials and decision criteria.

In my opinion, the work by Hsu et al. (2015) is a good example of how a work well written and structured that aims to report all the relevant aspects, can simultaneously be lacking key elements from a replicability perspective. The poor availability of materials and experimental procedures stood out from the beginning. For example, the key independent variable, the power-inducing music choice is only synthetically described. The original list with the 31 songs used in the pretest of the study by Hsu et al. (2015) is not comprised in the publication and I failed at getting an answer from the corresponding author. The original paradigm mentions a rationale behind the choice of songs from several musical genres (rock, hip-hop, sport music, ...) but does not provide any open access database or a list of the pretested stimuli, nor mentions a guiding principle to refer to.

This can be interpreted as an indicator of the quality of the original report, as suggested by Wicherts, Bakker and Molenaar (2011). In their article they found

evidence consistent with the idea that the willingness of authors to provide the collected data decreased when the publications presented reporting errors (such as imprecise p-values) and weaker statistical evidence against the Null Hypothesis (lower strength of the evidence). They analyzed a sample of publications from different psychological journals and sent all the corresponding authors a request to share their data. Afterwards, they proceeded to test the internal consistency of the findings of the paper, as well as the correctness of the reported statistics and p-values. This procedure allowed Wicherts and colleagues (2011) to draw the conclusion that especially when the reanalysis will more probably lead to disconfirming the original results, it is usually harder to get access to the original materials, data and details to verify them.

I think also the work by Hsu and colleagues (2015) might fall under these circumstances, whereby not clearly stating the logic behind their music choice and not being available to share it when contacted, reduces the faith we can have in the validity of the results, combined with the fact that my replication also failed to obtain them.

For what concerns my personal conclusions, with this thesis project I gained a new perspective on the work in psychological research, a valuable resource for my future career and studies, ideally in the field of academia. I had the opportunity to experience directly the steps behind an experiment, with the recruiting and administering process, as well as an occasion to apply my knowledge in the field of statistics guided by more qualified experts. Throughout my Bachelor's studies I had the perception that more and more courses and professors highlighted the importance of the replicability crisis, covering its historical background, the replication efforts, as well as the related issues and possible solutions. Hopefully, projects like mine, together with the awareness students gain through the open debate of this topic, will pave the way to future generations of professionals, who will create a more reliable psychological science, seeking to correct past mistakes and build stronger foundations for the whole field.

7. References

- * Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244. DOI: <https://doi.org/10.1037/0022-3514.71.2.230>
- * Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.
- * Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100, 407–425. DOI: <https://doi.org/10.1037/a0021524>
- Carney, D. R., Cuddy, A. J., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological science*, 21(10), 1363-1368.
- Carney, D. R., Cuddy, A. J., & Yap, A. J. (2015). Review and summary of research on the embodied effects of expansive (vs. contractive) nonverbal displays. *Psychological science*, 26(5), 657-663.
- Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J., ... & Schulte-Mecklenbeck, M. (2018). 7 Easy Steps to Open Science: An Annotated Reading List.
- Cuddy, A. J. C., Schultz, S. J., & Fosse, N. E. (2018). P-Curving a More Comprehensive Body of Research on Postural Feedback Reveals Clear Evidential Value for Power-Posing Effects: Reply to Simmons and Simonsohn (2017). *Psychological Science*, 29(4), 656–666.
- * De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological bulletin*, 127(6), 853.
- Fast, N. J., Gruenfeld, D. H., Sivanathan, N., & Galinsky, A. D. (2009). Illusory control: A generative force behind power's far-reaching effects. *Psychological Science*, 20(4), 502-508.

Garrison, K. E., Tang, D., & Schmeichel, B. J. (2016). Embodying power: A preregistered replication and extension of the power pose effect. *Social Psychological and Personality Science*, 7(7), 623-630.

Hsu, D. Y., Huang, L., Nordgren, L. F., Rucker, D. D., & Galinsky, A. D. (2015). The music of power: perceptual and behavioral consequences of powerful music. *Social Psychological and Personality Science*, 6(1), 75-83.

* Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. Proceedings of the National Academy of Sciences, 113, E4933–E4934. <http://dx.doi.org/10.1073/pnas.1608676113>

* Magee, J. C., Galinsky, A. D., & Gruenfeld, D. H. (2007). Power, propensity to negotiate, and moving first in competitive interactions. *Personality and Social Psychology Bulletin*, 33(2), 200-212.

* McGuire, W. J. (1983). A contextualist theory of knowledge: Its implications for innovation and reform in psychological research. In *Advances in experimental social psychology* (Vol. 16, pp. 1-47). Academic Press.

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, 7(6), 657-660.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, 7(6), 531-536.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on psychological science*, 7(6), 528-530.

Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S., & Weber, R. A. (2015). Assessing the Robustness of Power Posing: No Effect on Hormones and Risk Tolerance in a Large Sample of Men and Women. *Psychological Science*, 26(5), 653–656. <https://doi.org/10.1177/0956797614553946>

Renkewitz, F., & Heene, M. (2019). The Replication Crisis and Open Science in Psychology. *Zeitschrift für Psychologie*.

* Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 86, 638–641.

Schimmack, U. (2020). A meta-psychological perspective on the decade of replication failures in social psychology. *Canadian Psychology/Psychologie canadienne*, 61(4), 364.

* Schmidt, S. (2016). Shall we really do it again? The powerful concept of replication is neglected in the social sciences.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 22(11), 1359-1366.

Simmons, J. P., & Simonsohn, U. (2017). Power posing: P-curving the evidence. *Psychological science*.

* Simonsohn, U., Nelson, L. D., Simmons, J. P. (2014). P-curve: A key to the file drawer. *Journal of Experimental Psychology: General*, 143, 534–547.

* Smith, P. K., & Trope, Y. (2006). You focus on the forest when you're in charge of the trees: power priming and abstract information processing. *Journal of personality and social psychology*, 90(4), 578.

* Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on psychological science*, 7(6), 670-688.

Świątkowski, W., & Dompnier, B. (2017). Replicability crisis in social psychology: Looking at the past to find new pathways for the future. *International Review of Social Psychology*, 30(1), 111-124.

* Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences USA*, 113, 6454 – 6459.
<http://dx.doi.org/10.1073/pnas.1521897113>

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PloS one*, 6(11), e26828.

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, 39(4), 202.

*= *not directly consulted publications*

Acknowledgements

I would like to thank the statistical advice service of the University of Padova and Lucia Ronconi in particular, for the helpful guide and assistance provided for the data analysis and interpretation.

I would also like to thank all the participants who chose to take part to my pretests or to my experiment.

Appendix

Pre-registration module on Aspredicted.org

<https://aspredicted.org/8y9c8.pdf>

Songs used for the pretests

a. Complete list of songs in Spotify's playlist Top Hits Italia (02/02/2022)

1. La coda del diavolo – Rkomi, Elodie
2. Solite Pare (feat. the Supreme) – Sick Luke, Sfera Ebbasta
3. Blauer – Paky
4. Finchè Non Mi Seppelliscono – BLANCO
5. Pastello Bianco – Pinguini Tattici Nucleari
6. ∞ LOVE (feat. Guè) – Marracash, Guè
7. CRAZY LOVE – Marracash
8. MI FAI IMPAZZIRE – BLANCO, Sfera Ebbasta
9. SUI MURI – PSICOLOGI
10. abcdefu – GAYLE
11. Veleno – Guè
12. LAPROVINCE #1 – Rhove
13. KUMKITE – Salmo
14. Come nelle canzoni – Coez
15. SAPORE – with Tedua – Fedez, Tedua
16. Notti In Bianco – BLANCO
17. NEMESI (feat. BLANCO) – Marracash, BLANCO
18. Angelina Jolie – Bresh, SHUNE
19. DREAM TEAM (feat. Pyrex, Capo Plaza, Tedua & Shiva)
20. ME STAJE APPENNENN' AMÒ – LIBERATO
21. Heat Waves – Glass Animals
22. NUOVO RANGE (con SFERA EBBASTA) – Rkomi, Sfera Ebbasta, Junior K
23. La più bella – Mecna, CoCo
24. Enemy (with JID) – from the series Arcane League of Legends – Imagine Dragons, JID
25. Mi Fiderò (feat, Madame) – Marco Mengoni, Madame

26. Shakerando – Rhove
27. Paraocchi – BLANCO
28. TU MI HAI CAPITO (feat. Sfera Ebbasta) – Madame, Sfera Ebbasta
29. Pepas – Farruko
30. PARTIRE DA TE - Rkomi
31. FALENA (feat. Franco126, Coez & Ketama126) – Sick Luke
32. INDUSTRY BABY (feat. Jack Harlow) – Lil Nas X, Jack Harlow
33. TRAVESURAS (feat. MV Killa, Yung Snapp, Lele Blade, Vale Limbo) – SLF
34. Cold Heart – PNAU Remix – Elton John, Dua Lipa, PNAU
35. STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber
36. LONDRA – Rosa Chemical, Rkomi
37. LA CANZONE NOSTRA (con BLANCO) – MACE, BLANCO, Salmo
38. Antipatico – Sacky
39. Infinity – Jaymes Young
40. CHINGA (feat. Simba La Rue) – Rondodasosa, Nko, Simba La Rue
41. uNa DiReZioNe giUsTa – youngest Moonstar, Neffa, tha Supreme
42. Do It To It – ACRAZE, Cherish
43. Piango Sulla Lambo (feat. Rose Villain) – Guè, Rose Villain
44. IL GIORNO PIÙ TRISTE (feat. Ariete & Mecna) – Sick Luke, ARIETE, Mecna
45. CRY LATER (feat. Sfera Ebbasta & Luchè) – Noyz Narcos
46. Quanto ti vorrei – chiello, Shablo
47. Quello che fa male – LDA
48. Easy on Me – Adele
49. Scrivile Scemo - Pinguini Tattici Nucleari
50. perso nel buio (con Madame) – sangiovanni, Madame

b. List of the 31 songs pretested

1. La coda del diavolo – Rkomi, Elodie
2. Solite Pare (feat. the Supreme) – Sick Luke, Sfera Ebbasta
3. Blauer – Paky
4. Finchè Non Mi Seppelliscono – BLANCO

5. Pastello Bianco – Pinguini Tattici Nucleari
6. ∞ LOVE (feat. Guè) – Marracash, Guè
7. SUI MURI – PSICOLOGI
8. abcdefu – GAYLE
9. Veleno – Guè
10. LAPROVINCE #1 – Rhove
11. KUMITE – Salmo
12. Come nelle canzoni – Coez
13. SAPORE – with Tedua – Fedez, Tedua
14. Angelina Jolie – Bresh, SHUNE
15. ME STAJE APPENNENN' AMÒ – LIBERATO
16. Heat Waves – Glass Animals
17. La più bella – Mecna, CoCo
18. Enemy (with JID) – from the series Arcane League of Legends – Imagine Dragons, JID
19. Mi Fiderò (feat, Madame) – Marco Mengoni, Madame
20. TU MI HAI CAPITO (feat. Sfera Ebbasta) – Madame, Sfera Ebbasta
21. Pepas – Farruko
22. INDUSTRY BABY (feat. Jack Harlow) – Lil Nas X, Jack Harlow
23. TRAVESURAS (feat. MV Killa, Yung Snapp, Lele Blade, Vale Limbo) – SLF
24. Cold Heart – PNAU Remix – Elton John, Dua Lipa, PNAU
25. STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber
26. LONDRA – Rosa Chemical, Rkomi
27. LA CANZONE NOSTRA (con BLANCO) – MACE, BLANCO, Salmo
28. Antipatico – Sacky
29. Infinity – Jaymes Young
30. uNa DiReZioNe giUsTa – youngest Moonstar, Neffa, tha Supreme
31. Do It To It – ACRAZE, Cherish

Results of the pretests

a. SONGS – pretest (corrected values)

Mean of the evaluations of each song, derived from the average points received by them in each of the three dimensions of “Powerfulness”, “Dominance” and “Determination”.

Values are reported in descending order to be able to identify the three highest rated and three lowest rated pieces, respectively defining the high-power and the low-power conditions.

Reported in red are the two song pieces erroneously used as the third element for the experimental manipulation in both conditions, due to an error in the calculation spotted only upon later controls. “*abcdefu – GAYLE*” was used instead of “*LONDRA - Rosa Chemical, Rkomi*” for the high-power condition, while “*Angelina Jolie – Bresh, SHUNE*” took the place of “*SAPORE – with Tedua – Fedez, Tedua*” for the low-power one.

Music pieces	Mean	Standard Deviation
LA PROVINCE #1 - Rhove	5.45	1.41
La coda del diavolo - Rkomi, Elodie	5.07	1.20
LONDRA - Rosa Chemical, Rkomi	4.94	1.54
LA CANZONE NOSTRA (con BLANCO) – MACE, BLANCO, Salmo	4.67	1.67
Solite Pare (feat. the Supreme) – Sick Luke, Sfera Ebbasta	4.64	1.74
abcdefu - GAYLE	4.33	1.99
INDUSTRY BABY (feat. Jack Harlow) – Lil Nas X, Jack Harlow	4.01	1.33
Blauer – Paky	3.88	1.35
Do It To It – ACRAZE, Cherish	3.84	1.49
La più bella – Mecna, CoCo	3.81	1.73
ME STAJE APPENNENN' AMO - LIBERATO	3.81	1.36
Veleno - Guè	3.78	1.72

Pastello Bianco – Pinguini Tattici Nucleari	3.68	1.57
Enemy (with JID) – from the series Arcane League of Legends – Imagine Dragons, JID	3.64	1.54
Cold Heart – PNAU Remix – Elton John, Dua Lipa, PNAU	3.48	1.68
Infinity - Jaymes Young	3.44	1.31
SUI MURI - PSICOLOGI	3.43	1.51
KUMITE – Salmo	3.42	1.46
Come nelle canzoni - Coez	3.41	1.39
uNa DiReZioNe giUsTa – youngest Moonstar, Neffa, tha Supreme	3.41	1.79
Finchè Non Mi Seppelliscono – BLANCO	3.32	1.67
Angelina Jolie – Bresh, SHUNE	3.29	1.34
Mi Fiderò (feat, Madame) – Marco Mengoni, Madame	3.15	1.76
TU MI HAI CAPITO (feat. Sfera Ebbasta) – Madame, Sfera Ebbasta	3.11	1.77
TRAVESURAS (feat. MV Killa, Yung Snapp, Lele Blade, Vale Limbo) – SLF	3.07	1.43
Heat Waves – Glass Animals	2.77	1.10
Pepas - Farruko	2.68	1.45
Antipatico - Sacky	2.65	1.30
SAPORE – with Tedua – Fedez, Tedua	2.61	1.15
STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber	2.58	1.35
∞ LOVE (feat. Guè) – Marracash, Guè	2.57	1.34

b. LYRICS – pretest

Mean of the evaluations for each song, derived from the average points received by them in each of the three dimensions of “Powerfulness”, “Dominance” and “Determination”.

Music pieces	Mean	Standard Deviation
Finchè Non Mi Seppelliscono – BLANCO	3.92	1.81
Enemy (with JID) – from the series Arcane League of Legends – Imagine Dragons, JID	3.85	1.77
abcdefu - GAYLE	3.69	2.05
LA CANZONE NOSTRA (con BLANCO) – MACE, BLANCO, Salmo	3.42	1.48
La coda del diavolo – Rkomi, Elodie	3.42	1.60
INDUSTRY BABY (feat. Jack Harlow) – Lil Nas X, Jack Harlow	3.41	1.99
Infinity - Jaymes Young	3.22	1.59
Mi Fiderò (feat, Madame) – Marco Mengoni, Madame	3.14	1.71
Pepas - Farruko	3.02	1.82
SAPORE – with Tedua – Fedez, Tedua	2.96	1.42
∞ LOVE (feat. Guè) – Marracash, Guè	2.94	1.31
KUMITE – Salmo	2.91	1.38
Come nelle canzoni - Coez	2.90	1.44
Do It To It – ACRAZE, Cherish	2.86	1.78
Heat Waves – Glass Animals	2.84	1.55
SUI MURI - PSICOLOGI	2.84	1.49
uNa DiReZioNe giUsTa – youngest Moonstar, Neffa, tha Supreme	2.78	1.55
TU MI HAI CAPITO (feat. Sfera Ebbasta) – Madame, Sfera Ebbasta	2.73	1.57
Veleno - Guè	2.66	1.29
Pastello Bianco – Pinguini Tattici Nucleari	2.60	1.38

STAY (with Justin Bieber) – The Kid LAROI, Justin Bieber	2.59	1.39
Cold Heart – PNAU Remix – Elton John, Dua Lipa, PNAU	2.57	1.33
La più bella – Mecna, CoCo	2.56	1.18
Angelina Jolie – Bresh, SHUNE	2.55	1.48
Antipatico - Sacky	2.50	1.75
TRAVESURAS (feat. MV Killa, Yung Snapp, Lele Blade, Vale Limbo) – SLF	2.45	1.60
Blauer – Paky	2.41	1.64
LONDRA - Rosa Chemical, Rkomi	2.34	1.27
Solite Pare (feat. the Supreme) – Sick Luke, Sfera Ebbasta	2.17	1.47
LA PROVINCE #1 - Rhove	2.06	1.51
ME STAJE APPENNENN' AMO - LIBERATO	1.62	1.15

Die Rolling Task

*Ti chiedo ora di immaginare il seguente scenario:
 Puoi prendere parte a un gioco, per cui se riuscirai a prevedere correttamente
 l'esito del lancio di un dado a sei facce potresti vincere 5 euro.
 Puoi scegliere di tirare il dado tu stesso o di lasciare che lo faccia lo
 sperimentatore al posto tuo.
 Cosa faresti?*

- Sceglierei di tirare io il dado
- Sceglierei di fare tirare il dado allo sperimentatore