An Experimental Psychological Study

The Effect of Music on Performance of a Task

Abstract
In order to investigate whether music affected performance of a task, and experimental technique was used, variables were manipulated and data recorded.

The aim of this study was to investigate whether different music styles affected the performance of a task. It was a novel experiment, only loosely based on previous research dating back to the 19th century.

The method involved three groups of participants undertaking a test (solving thirty anagrams). One group had fast music in the background, one had slow music and the third performed it in silence. The participants were primarily selected via a systematic sample, but this would have been changed to an opportunity sample had some participants not turned up.

It was hypothesised that there would be significant differences between a) fast and slow music, b) fast and no music and c) slow and no music.

A two-tailed Mann-Whitney U test at a significance level of p=0.05 revealed that all three alternative hypotheses were accepted and null hypotheses rejected.

The data collected illustrated that having slow music playing in the background improved performance of the task compared to performing it in silence, while fast music worsened performance.

The implications of this study, its limitations and suggestions of follow up studies will be further discussed.

Contents

Introduction

Social influence describes how other people around us can influence our actions. It is especially relevant in situations where groups of people are performing a task together, as discovered by Triplett in one of the first social influence experiments, conducted in 1898. He found that when children were asked to spin a fishing reel, it was spun faster when they were performing in groups than on their own. This effect was termed 'social facilitation', as the presence of others appeared to help, or facilitate, the person performing the task. This idea was also supported by Allport (1920), who demonstrated that college students performing multiplication tests also worked faster alongside other students. But there is an opposite side to social facilitation - social loafing. This describes the process whereby when in a group, an individual puts less effort into a task. Latane, Williams and Harkins (1979) demonstrated this to good effect when they found that when children were asked to be noisy, they were more quiet when they were in the group and were louder on their own.

Another factor that could affect performance of a task is music. From Beethoven to the Bee Gees, music has had an impact on most of us somewhere in our lifetime. It is only comparatively recently however that it has become seriously analysed and tested in different situations in order to recognise its true effects. Many of us have at some time associated a song or piece of music with an event or situation in our lives, whether good or bad, and on hearing it again can recall feelings and emotions of that situation. It is also deeply representative of individual cultures and by listening to the music of a certain culture, we can learn a great deal about it. Its diversity is huge - the Internet itself has nearly 700,000 web pages dedicated to it - as are its range of uses - music therapy is now being used by psychiatrists to relax and help patients. With such a broad topic to consider there were many possible paths to take, but the avenue that seemed most interesting to explore was that if different types of music would affect the performance of a task. Being a student, I was especially interested to see what sort of music affects work quality, with a view to using it personally with school-work and revision. I have found that when studying, slow relaxing music helps me concentrate and produce a better quality of work, while other students have also reported being able to revise better with slow, relaxing music playing in the background. Of the studies I had access to, none provided any clear cut results in this area, although Wallace (1994) found that a simple repetitive melody increased memory of text. According to Turner et al. (1996), the amplitude of a melody is important for task performance. They discovered that lower amplitude elicited a

slower response time to unexpected visual events.
The basis of this study is a similar one undertaken by Mayfield and Moss (1989), who asked students to perform a task with fast, then slow tempo music playing in the background. The students reported afterwards that the fast music acted as a stress stimulator, although it made them perform the task quicker. In accordance, the participants reported that slower music seemed to relax them, thus slowing down the speed of their work. My study however focuses on the quality of the work rather than the speed at which it was performed, where a study by Stough et al. (1994) becomes relevant. They performed IQ tests on participants, with one group having Mozart playing in the background and the other performing it in silence. Here, the slow relaxing music had no effect on the results, arguing the case that music has no effect on test performance.

Therefore, based on the above research this study aims to test whether music affects performance of a task. Also due to the lack of conclusive research, my hypotheses are non-directional.

Hypotheses

1)
H1: There will be a significant difference between the scores participants get on the anagram test depending on whether they were in the condition where fast or slow music was played.
H0: There will be no significant difference between the scores participants get on the test, despite whether they were in the condition where fast or slow music was played.

2) H1: There will be a significant difference between the scores participants get on the anagram test depending on whether they were in the condition where fast music or none at all was played.
H0: There will be no significant difference between the scores participants get on the anagram test, despite whether they were in the condition where fast music or none at all was played.

3) H1: There will be a significant difference between the scores participants get on the anagram test depending on whether they were in the condition where slow music or none at all was played.
H0: There will be no significant difference between the scores participants get on the anagram test, despite whether they were in the condition where slow music or none at all was played.

Method

Design:

An experimental procedure, in particular a lab experiment, was used in which all other variables apart from the independent and the dependant were fixed. An independent samples design was selected for the participants to rule out any order effects.

Performance was tested by using a sheet of thirty anagrams. These anagrams were between four and ten letters long and consisted of randomly selected every-day words. There was a mix of short and long words (listed in Appendix I) allowing a wide range of results to analyse and a means of easily gauging any differences in the groups. Anagrams were chosen as they were very frustrating, as one participant said, "...because the answer is right there, but it's also the question!" It was also ensured that the words selected did not form two different words. There was also an extra incentive to perform well - the participant who had the most correct answers won a prize - a bag of sweets was used for this particular experiment. This was designed to again raise the competition levels during the task.

A time limit of five minutes was decided upon in which to solve the anagrams. The independent variable was the music. One group had slow classical music playing in the background and another had fast dance music. A control group was used in order to compare the results of those who performed the task with music to a group with no independent variables to affect their performance. Both pieces of music were played at the same volume, -54 decibels. The dependant variable was the score on the anagram test.

To identify any confounding variables and test the general running of the experiment, a pilot study was carried out on five participants of the same age group and using the same sampling methods as those on the real study, with no music playing in the background. The main objective was to find a suitable time limit in which the task was to be attempted. The original time limit was seven minutes, but on completion, the participants reported that a shorter time limit would offer a higher level of competition. There were no confounding variables identified.

To test for significance, a Mann-Whitney U test was used at a significance level of $p = 0.05$, with an individual test being performed for each pair of results (fast music/slow music, fast music/no music, slow music/no music). The justification for this is in the results section.

Participants:

The target population was the Sixth form of Arnewood School, Hampshire. These students were aged 17/18 and of mixed sex. First, thirty participants were selected by systematic sampling, selecting every tenth person on a register of the Sixth Form until thirty had been selected. These were then randomly allocated into their experimental groups of ten by the researcher, via manual selection - names drawn out of a large box, but combining this with an opportunity sample. This also produced a representative sample of the academic side of the Sixth form with every subject represented in the participants, improving the validity of the experiment.

Materials:

Materials used consisted of participants' question and response sheets (Appendix I), an answer sheet (Appendix II), standardised instructions (Appendix

III), a bag of sweets (prize), a Sony stereo, a stopwatch and tracks from two CD's (detailed below).

'Slow' music:
J.S. BACH, Suite No.3, 'Air', from Essential Classics CD, Deutsche Grammophon records (1990).

'Fast' Music:
From 'Trance Nation' CD, Ministry Of Sound Records, (2000)

Procedure:

The randomly assigned groups of ten participants arrived separately at a classroom at the appointed time. Here the standardised instructions, listed in Appendix III, were read to all participants. The anagram/response sheets (one allocated to each participant) were already lying face down in front of the seating positions accompanied by a piece of writing equipment. The participants were asked to be seated and make themselves as comfortable as possible, before the standardised instructions were told to them. The participants were then told to ensure their names were not to be written anywhere on the paper for ethical reasons, and were then instructed to start. Here, the music for the two groups that required it and the stopwatch were started. After five minutes the music was stopped and the participants were instructed to stop writing and hand their results to the researcher who then added up the results and issued the prize to the winner. Throughout the experiment, participants were not allowed to confer and could only talk once all the papers had been handed in. The participants were then debriefed, shown the answer sheet, and free to leave, after the researcher had made it clear that none of the other participants yet to be tested be told of the aims of the experiment or what it entailed.

Results

Hypothesis One

The following table illustrates the differences in results between the groups who had slow music and fast music in the background:

Table 1: Score comparison between fast and slow music

Table one shows that there was an evident difference between performance of the two groups, suggesting that having slow music playing in the background enhances performance, compared to fast music.

In order to test the significance of the results, a Mann Whitney U test was performed. This test was chosen because of its suitability as a means of measuring difference between groups of results at an ordinal and unrelated level. The significance level of 0.05 (5%) was selected as it is accurate enough to give a good certainty of correctness, but is not too precise as to be excessive. A two-tailed test

was decided upon as the background research proved to be inconclusive as to the expected outcome.

Statistical Results:

The Mann Whitney U test can be found in Appendix V.

The calculated value of U was 30, which is greater than the critical value of 23. This shows that there was a significant difference between the fast and slow music conditions. Because of this, the null hypothesis is rejected.

The raw results for all hypothesis can be found in Appendix IV.

Hypothesis Two

The following table illustrates the differences in results between the groups who had fast music and no music in the background:

Table 2: Score comparison between fast music and none

Table two shows that there was a small difference between performance of the two groups, suggesting that having no music playing in the background at all slightly enhances performance of a task, compared to having fast music playing.

In order to test the significance of the results, a Mann Whitney U test was performed. The justification for this test was outlined in the results for hypothesis one. Again, a two-tailed test was decided upon as the background research proved to be inconclusive as to the expected outcome.

Statistical Results:

The Mann Whitney U test can be found in Appendix VI.

The calculated value of U was 56.5, which is greater than the critical value of 23. This shows that there was a significant difference between the fast and no music conditions. Because of this, the null hypothesis is rejected.

Hypothesis Three

The following table illustrates the differences in results between the groups who had slow music playing in the background and none at all:

Table 3: Score comparison between slow music and none

Table three shows that there was a small difference, equal to one correct answer per participant, between performance of the two groups. This suggests that having slow

music playing in the background slightly enhances performance of a task, compared to having no music playing at all.

In order to test the significance of the results, a Mann Whitney U test was performed. The justification for this test was outlined in the results for hypothesis one. Again, a two-tailed test was decided upon as the background research proved to be inconclusive as to the expected outcome.

Statistical Results:

The Mann Whitney U test can be found in Appendix VII.

The calculated value of U was 39.6, which is greater than the critical value of 23. This shows that there was a significant difference between the fast and slow music conditions. Because of this, the null hypothesis is rejected.

Discussion

The aim of this investigation was to see what effect music has on a task. The results show that different kinds of music have a different effect on the listener. Slow music increased the score on the anagram test and fast music had a detrimental effect when compared to having no music at all.

In hypothesis one, the effect of the two types of music was evident. In fact, the difference in test performance between the two variables was the largest of the three sets of results. This clearly demonstrates the effect of the music styles on the participants - table one shows that those listening to the slow music had a mean average of almost two points each more than those listening to fast music. This was a surprising fact, as previous research did not suggest much of a difference, if any at all, between the two types of music. The alternate hypothesis, that there will be a significant difference in the score on the anagram test depending whether fast or slow music is played, is accepted and the null hypothesis, that there will be no significant difference, is rejected.

In hypothesis two, it was demonstrated that not only did fast music not improve performance, but in fact it had a detrimental effect on the participants. As table two shows, those performing the task with no music performed better than those with fast music in the background. It also shows that those listening to fast music had a mean average of one point each less. Also, this hypothesis had the largest difference between the observed and critical values of U. Despite this, the difference between the total scores was the smallest of the three hypotheses. The alternate hypothesis, that there will be a significant difference in the score on the anagram test depending whether fast music or none at all is played, is accepted and the null hypothesis, that there will be no significant difference, is rejected.

Hypothesis three was effectively the test to see which of the two music styles produced the best performance. Table three shows that slow music improved performance by a mean average of one point per participant. This revealed that slow music improved performance on the test more than the other two music styles. The alternate hypothesis, that there will be a significant difference in the score on the

anagram test depending whether slow music or none at all is played, is accepted and the null hypothesis, that there will be no significant difference, is rejected.

Not only were the mean and total scores close together, they were also very low. Considering all scores were out of thirty, scores between eight and twelve were the most common of all three groups. This may have been due to the competition and time limit putting pressure on the participants.

Graph one on page twelve shows the differences between the mean average test score for the three music styles. The performance with fast music playing falls below that with none at all and even further behind that of slow music which stands visibly ahead with the highest mean average score.

This study was not based on one single piece of previous research, but many associated studies. For this reason, my results cannot be directly compared with those of another study, although they can be linked to the study by Mayfield and Moss (1989). They found that fast music acted as a stressor to participants while slow music relaxed them, allowing them to perform the task quicker. When debriefing the participants, five that took the test with the slow music commented that they felt relaxed and believed it was because of the music. However, the group with the fast music did not believe it affected their performance. Interestingly, when asked whether they tried harder because of the presence of other people in search of the prize, four participants in the fast music group commented that they had, while no participants in any other group did. This may indicate that the fast music stimulates social facilitation, an addition to the studies by Triplett (1898) and Allport (1920). However, the low scores themselves may indicate the direct opposite to this; social loafing. This idea suggests that the presence of others in a group performing the same task elicits a poorer performance. However, to test this properly the participants must be tested individually to compare the scores, a study that would require concentration on social influence. The results of my study contrast those of Stough et al.(1994), who found that classical music did not affect the performance of an IQ test. However, an IQ test sets questions which the participant may not know the answer to no matter what music is played, whereas my study involved questions where the answers were constantly present, the music simply relaxed/stressed the participants thus affecting how they coped with the anagrams.

One of the limitations of using only an anagram test was the unknown vocabulary of the participants - some may have only known the short words and never come across the longer ones. This was where the cross section of subjects became important (English students were present in every group). It was for this reason many short words were chosen, of which every participant correctly worked out nearly all of them. A major drawback with this type of study is that a researcher can never be sure whether they are measuring the influence of the music or the purely the intelligence of the participant (whether the music had any impact). The only way of being certain of this is to use the same students in the same test with the different music styles, but which would of course be pointless as after the first test they would know most of the answers. Therefore, the most accurate way of doing this is to use students close in intelligence, age etc.

Because of the size of the groups due to time constraints, I expected a very small difference in scores thus making analysis difficult. However, the results proved to be varied enough to notice the trends and I believe that larger groups would only emphasise these trends. Another limitation concerning the participants is the process of selection. An opportunity sample was the final selection process after the participants had been selected via systematic sampling, when a more representative sample may have offered a wider range of results. One major limitation is generalisation. It would be very difficult to generalise the findings of my study to other age groups or other tasks, but the results could be useful in relation to other studies on this subject. I had anticipated possible problems with availability of participants, but had a full compliment for all three experiments, thus eliminating any problems on this front. I would have preferred a more accurate way of sampling than combining random with opportunity, but as I received a full compliment of participants for all three experiments this is less of a consideration. Testing a number of bigger groups, about twenty per group, would have allowed me to use a number of 'tests' (anagrams, crosswords, maths questions) to examine whether it was indeed the music affecting performance, and not the difficulty or type of task (English style questions, maths style questions etc.).Overall, I was very pleased with the set up and running of the experiment. Because the experiment was carried out in an unnatural environment, ecological validity is also a limitation of the study. The best way to solve this is to incorporate the selected task, with music when necessary, into a normal school lesson as this represents a more natural environment for the participant.

This study will impact students and, in fact, anyone looking for ways to increase their concentration on work or quality of work etc. It was aimed at students like myself- to make aware the potential of music as an aid to revision/work. However, a study like this will be totally disregarded by someone who only likes and only works to one specific music type. My study simply suggests that a slower type of music may improve performance of a task, of which homework could be one. With more research on different music types, different age groups and different tasks, it could be ascertained as to what types of music affect what types of task etc., but without this further research my study cannot stand alone and be generalised to these other areas. This study could be seen as the first step into investigating the subject of music aiding performance. With more published research in this area, the effect of music could be made more publicly aware. As music is such a huge industry with an even bigger following, people could learn about the types of music that would help them possibly improve their work though their interest for music. Ideally, my study would act as this stepping stone to further research.

There are many directions to go with follow up studies, whether focusing on music or not. A possible future study following on from this could focus on social facilitation as an influencing factor. Future researchers could perform the same test but remove influencing factors and compare the results. For example, if the prize was removed you may expect the level of competition to go down; if the time limit was removed you may expect the pressure on participants to go down, thus yielding higher mean results. This may also help to identify the extent to which social facilitation/loafing affects the performance of the task. Alternatively they could investigate the music aspect, trying many different types of music (rap, dance, classical etc.) and seeing which genres of music give better results. A study involving personal preferences of music of participants could also be quite useful. The researcher could ask beforehand in a questionnaire the favourite music type of the participant. They could then perform a task listening to that music, and then perform another task listening to

their least favourite music style and compare the results. This would indicate fully the amount to which music affects the performance of a task. It also would be sensible to undertake different types of task (maths, English etc.) to allow valid generalisation, in accordance with different age groups. For students, researchers could associate different types of music with revision for a test and compare performances. This may allow students to see which types of music may be useful for them in revision for exams. However, this may still depend heavily on personal preference.

Conclusion

Overall, the results of my study show that slow (classical) music can improve the performance of a task and fast (dance) music can worsen the performance, compared to none at all. It also suggests that fast music may encourage social facilitation.

Appendices

Appendix I - Participant Question/Response Sheet

sowhad -
konmey -
flufcitid -
wilrings -
noteaudic -
stainnoum -
sacslic -
nolcyba -
zigamane -
ningnigeb -
dicsedover -
dubliing -
cagedroune -
cadeox -
mipsle -
krow -
henop -
divoe -
pice -
rakcc -
tird -
kilm -
korf -
kalt -
wolb -
perap -
tighl -
dinwow -
heset -

verco -

Appendix II - Answer Sheet

sowhad - shadow
konmey - monkey
flufcitid - difficult
wilrings - swirling
noteaudic - education
stainnoum - mountains
sacslic - classic
nolcyba - balcony
zigamane - magazine
ningnigeb - beginning
dicsedover - discovered
dubliing - building
cagedroune - encouraged
cadeox - coaxed
mipsle - simple
krow - work
henop - phone
divoe - video
pice - epic
rakcc - crack
tird - dirt
kilm - milk
korf - fork
kalt - talk
wolb - bowl
perap - paper
tighl - light
dinwow - window
heset - sheet
verco - cover

Appendix III - Standardised Instructions

All three groups were read this set of instructions:

? The task I would like you to perform is to work out a list of thirty anagrams. You have five minutes to work out as many as you can and there will be a prize for the person with the most correct answers.
? Is there anyone who feels uncomfortable with the task they have been asked to do?
? I will signal when you may start and after five minutes I shall ask you to stop and give me your answer sheets. I will request that from this point onwards there is no talking until all the papers have been handed in.
? If there is anyone who wishes to leave would they please indicate so now.
? (after the five minutes had elapsed) Could you please stop writing and hand in your sheets, making sure you have not written your name on the sheet.
? (after the sheets had been handed in, the results added up, the prize issued and the participants had seen th4e answers and been debriefed) Thank you for participating in this study. Please do not tell anyone about any details or aims of the study you have just undertaken.

Appendix IV - Raw Results
Appendix V - Mann Whitney U Test For Hypothesis One

$$U = 10 \times 10 + \frac{10\,(10 + 1)}{2} - 125 = 30 \quad U = 30$$

$$U' = 10 \times 10 - 30 = 70 \quad U' = 70$$

Observed value of U = 30

Critical value of U = 23 at a significance level of p?0.05 and a two-tailed test.
Therefore, as 30 > 23 the null hypothesis is rejected.
Appendix VI - Mann Whitney U Test For Hypothesis Two

$$U = 10 \times 10 + \frac{10\,(10 + 1)}{2} - 111.5 = 43.5 \quad U = 43.5$$

$$U' = 10 \times 10 - 43.5 = 56.5 \quad U' = 56.5$$

Observed value of U = 43.5

Critical value of U = 23 at a significance level of p?0.05 and a two-tailed test.
Therefore, as 43.5 > 23 the null hypothesis is rejected.
Appendix VII - Mann Whitney U Test For Hypothesis Three

$$U = 10 \times 10 + \frac{10\,(10 + 1)}{2} - 115.5 = 39.5 \quad U = 39.5$$

$$U' = 10 \times 10 - 39.5 = 60.5 \quad U' = 60.5$$

Observed value of U = 39.5

Critical value of U = 23 at a significance level of p?0.05 and a two-tailed test. Therefore, as 39.5 > 23 the null hypothesis is rejected.