Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Safe and Sound: An investigation of risk homeostasis and the influence of music preference and exposure in a video game environment.

Name: Doris van Duijl
Student number: s1125478

Date: 18-08-2017
Supervisor: Jop Groeneweg
Second reader: William Verschuur
Cognitive Psychology
Thesis Msci Applied Cognitive Psychology

Abstract

This study examines the Risk Homeostasis Theory as described by Wilde (1982). According to this theory, risk behavior is influenced by an individual's desired risk level and experienced risk level. When experienced risk is lower than desired risk (possibly due to protection), compensatory risk-taking is expected to occur in order to reunite these factors.

Additionally, the influence of music involvement and emotional mediation on risk-taking was considered. Previous research on the relation between musical preference and risk-taking has indicated that part of its process involves emotions. In this study, factors used for analysis are music preference, music exposure in hours and emotional susceptibility. It is expected that these factors have an influence on the level of participants' risk-taking.

The experiment was conducted in a controlled setting, with the aim to minimize methodological issues associated to the Risk Homeostasis Theory. 69 participants between the ages of 18 and 36 were presented with a questionnaire on music involvement and Doherty's (1997) Emotional Contagion scale. For risk homeostasis, a computer game called 'The Spaceship Game' was used. It consisted of 5 sessions in which participants flew a spaceship while avoiding a field of incoming meteors. In each session, between 0 and 5 shields were present for protection in case of a collision. These shields were the in-game equivalent of experienced risk. Risk-taking was measured through parameters speed, time to collision (TTC) and distance to the closest meteor (DCM). It was expected that more shields (and thus less experienced risk) would lead to compensatory risk-taking, expressed in high speed and low TTC and DCM.

Evidence for the occurrence of risk homeostasis was found, mainly on a large, between-sessions scale: when more shields were present, speed was higher and TTC and DCM were lower. The analysis of music involvement as a factor in risk-taking yielded little results by itself, but interesting findings emerged with emotion into the equation. Participants with high susceptibility to fear tended to fly more carefully, in some cases significantly so. In case of music preference, Electronic music lovers showed this difference inversely. From these results it can be concluded that risk homeostasis occurred in the game environment, but ample opportunity remains for further research. The music and emotion approach requires further investigation.

The results of this study add to the currently mixed body of research on risk homeostasis. The findings on music can provide new insights into music as a component of risk-taking, and the nuancing role of emotions. Current findings are a starting point for future research, and can be cautiously interpreted for specialized safety measures and training.

Key words: risk homeostasis, risk compensation, music preference, emotional contagion.

1. Introduction

June, 2017. Coffee manufacturer Douwe Egberts discontinues a campaign featuring a group of young adults euphorically jumping off a bridge only 2 weeks into its run. An article in De Volkskrant (van der Velden, 2017) describes how this attempt at promoting a new line of iced coffee raised government concerns about encouraging risky behavior, leading to its early demise. Whether adolescents far and wide would actually start plunging to their deaths from an advertorial image can be questioned, but risk-taking on a more common level is unquestionably a force to be reckoned with. In many fields such as traffic, work and health, taking too many risks can lead to serious accidents. Even so, people often choose the more dangerous option: A CBR report on driving behavior in The Netherlands (2017) shows that around 12,5% of drivers had a habit of sending text messages while driving, even though 98,7% deemed this (highly) dangerous.

In the case of texting while driving, an erroneous decision by a human would be the main cause of an accident. In light of earlier research, this isn't unusual. In a study by Lu (2006), human error is described as responsible for "nearly all" traffic accidents. This description is not new, as in 1987 Hale and Glendon (1987) mentioned an approach known as the '80:20 rule' as an often-used as the ratio of human error to technical error. Since humans are so error-prone, why do they take risks?

1.1 Factors in risk-taking

Since the presence of risks alone does not always stop people from taking them, other variables must play a part in risky behavior. In earlier stages of psychology, Freud (1904, as discussed in Hale & Glendon, 1987) viewed accidents as an individual's unconscious way of self-punishment for improper thoughts. Frustrated energy would build up and cause humans to damage themselves as a result. Many years and psychological findings later, research points to directions such as more willingness to take risks with a positive mood (Johnson & Tversky, 1983), a high self-esteem (Josephs, 1992) and a sensation-seeking personality (Horvath & Zuckerman, 1993). Utilitarian reasons like avoiding losses (Kahneman, 2003) and weighing gains (Fishhoff et al., 1987) influence the willingness to take risks as well. A combination of factors was found in a study by Isen and Patrick (1983), where a positive mood led people to take more risks, but only if the consequences would allow their happy mood to be maintained. Although risk is often described in terms of willingness, it has to be noted that not all risk is taken completely voluntarily. In their study on gain-related risk-taking, Fishhoff et al. (1987) noted that people were not always aware of the amount of risk they were in, for example industrial workers. If this risk is technically phrased or difficult to imagine, it can be underestimated, leading to a higher accident rate.

1.2 Risk Homeostasis Theory

The estimation of risk plays a large role in how the amount of risk-taking is determined, according to Wilde (1982) and his Risk Homeostasis Theory (RHT). In this theory, risk behavior is described as part of a constantly updating homeostatic feedback system. It describes that people are constantly adapting their risk-taking to the situation they are in. Two key factors in this theory are the **amount of perceived risk** and the **amount of desired risk**. Perceived risk is mostly determined by the environment, for example wearing a seatbelt or driving speed. Desired risk is an individual's preferred level of risk, determined by their personal traits and considerations of its costs and benefits (Wilde, 2014). These are established through four factors:

1 The benefits of risky behavior *3* The benefits of safe behavior
2 The costs of risky behavior *4* The costs of safe behavior

If the expected benefits of risky behavior outweigh those of safe behavior and/or the costs of safe behavior are higher than those of risky behavior, the desired amount of risk can be expected to go up. If the increasing desired risk transcends perceived risk, this results in risk-taking.

Wilde (1982, 2014) describes the desired risk level as being relatively constant, approximated by a fluctuating perceived risk level, dependent on the situation. When the difference between perceived risk and desired risk grows too wide, humans adapt their behavior to contain more or less risk to reach an equilibrium.

With driving behavior as an example, Wilde (1982) compares this process to that of a thermostat, as shown in figure 1. In this comparison, desired risk would be the set temperature on the thermostat dial, and perceived risk the measured room temperature. If the room temperature is lower than the set temperature, the system starts heating.

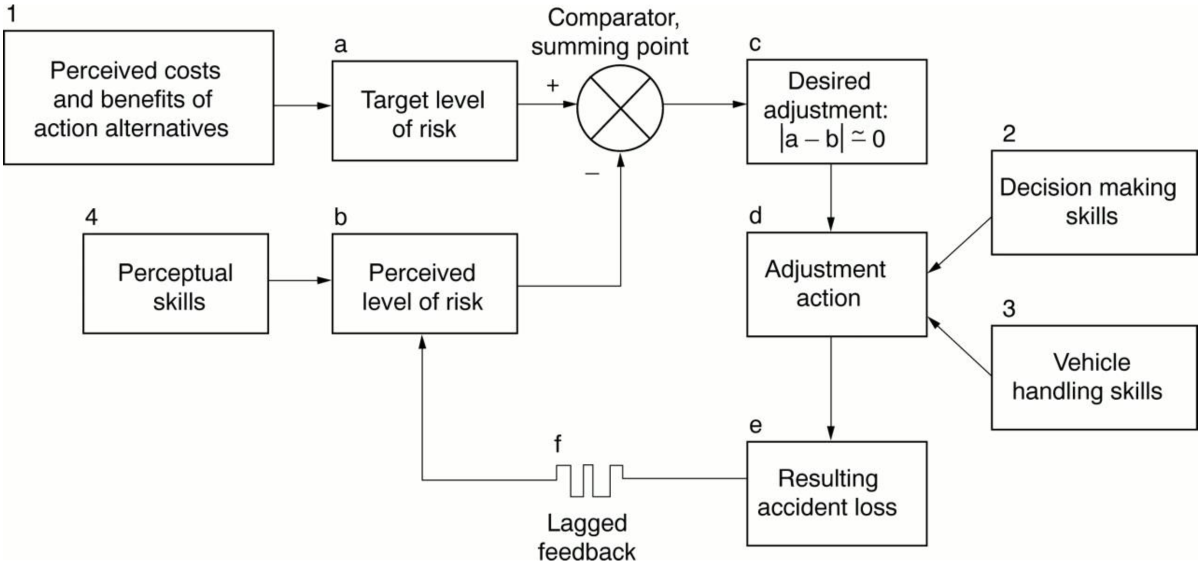

Figure 1. The risk homeostasis model, as proposed by Wilde (1982).

In traffic this would mean: if the amount of perceived risk while driving is lower than comfortable (for example, because of a seatbelt), behavior will shift to more reckless driving to move perceived risk back to the desired level. The same goes the other way around: if the situation becomes too dangerous, driving will become cautious to lower the amount of perceived risk. This behavior is determined by two factors: the individual's decision making and vehicle handling skills. These respectively influence the way compensatory behavior is decided on and how it unfolds in reality. The outcome of this homeostatic behavior adjustment, in turn, leads to a certain accident rate. The accident rate eventually reaches the individual, influencing their perceived amount of risk, and subsequently the desired 'target' point of risk. Because accident rate feedback is not apparent instantly, this process is called 'lagged feedback'.

As influencing the level of perceived risk eventually influences desired risk as well, the effects of preventive measures such as warning signs are of temporary nature psychologically. Physical safety measures like seatbelts and anti-lock brakes (ABS) can even result in dangerous driving through this same channel (Wilde, 2002). Thus, adjusting individuals' desired risk level instead would be a more fruitful approach (Wilde, 1982, Simonet & Wilde, 1997, Wilde, 1998). Wilde et al. (2002) concluded that a reward-based, target risk-focused approached was the most fruitful, such as rewarding drivers for safe behavior with discounts and bonuses instead of punishing unsafe behavior with fines. Schmidt (1987, as discussed in Wilde, 1989) argued that, even more than materialistic rewards, influencing people's intrinsic morals and awareness of risk could be the key to voluntary safe behavior.

1.3 Risk Homeostasis Theory: Controversy in research

Up until now, research on the Risk Homeostasis Theory has produced mixed results. Although evidence has been found in various aspects of life, issues with the theory are mainly methodological in nature (Hoyes & Glendon, 1993; Hoyes et al., 1996; Evans, 1987; Elvik, 2004; Adams, 1988; McKenna, 1987; O'Neill & Williams, 1998). Both sides will be discussed.

1.3.1 Research in favor of RHT

Although applied in various fields, the body of research on risk homeostasis in traffic settings is by far the most extensive. Studies vary from real-life data to driving simulators. Support from real-life accident data was found and discussed by Wilde (1998) and Wilde et al. (2002), after a law in Sweden changed driving directions from left to right. The implication of this change was followed by a decrease in car crashes, indicating that drivers' perceived amount of risk had gone up, followed by more careful behavior. Aschenbrenner and Biehl (1994, as discussed in Trimpop, 1996) found evidence of risk homeostasis in taxi drivers in München. Out of all their taxis, half were equipped with ABS and half without. Results showed that the taxi drivers with ABS drove faster than those without and performed more dangerous maneuvers. The same results were found by Grant and Smiley (1993), who described an increase in reckless behavior in cases of installed ABS. Assum et al. (1999) found evidence for risk homeostasis in an increase in drivers' speed and decrease of concentration after the installation of road lights on the Norwegian E18-route.

RHT support from simulated environments was obtained by Jackson & Blackman (1994). They approached risk-taking from a motivational viewpoint: influencing individuals' desired risk by manipulating the costs of an accident. This method bore fruit, as accident prevalence went down as a result of a higher cost. Hoyes, Stanton and Taylor (1996) followed up with their own driving simulator experiment, this time with a focus on perceived risk level. During the experiment, participants were presented with both a low-risk and high-risk environment to drive through. Changing conditions showed that significantly less risky behavior occurred in the high-risk environment. Although these results supported the existence of risk homeostasis, a contradictory finding was the speed with which participants adjusted. According to the Risk Homeostasis Theory feedback occurs slowly, but participants showed changes in behavior in as little as a 10-minute timespan. Hoyes et al. (1996) theorized that a possible explanation could be that changes in the risk environment were obviously recognizable to participants, leading them to rethink their strategy. The implications of driving simulator experiments to investigate risk homeostasis were discussed in studies by Hoyes, Dorn, Desmond and Taylor (1996) and Glendon, Hoyes, Haigney and Taylor (1996). Although these studies produced partial support for the notion of risk homeostasis, their writers emphasized the need for caution when generalizing simulator results to real-life situations.

Despite its mainly traffic-oriented approach, the Risk Homeostasis Theory has been applied in multiple fields. Hoyes (1994) discussed research on risk homeostasis in air-traffic control and nuclear power plant control, without finding clear evidence for the theory at that time. Despite these results, he argued that the theory needed proper consideration outside traffic context, as its mechanisms might not be universally applicable. Stetzer & Hoffmann (1996) discussed the implications of RHT for occupational contexts as well, for example the use of safety equipment. Their considerations were not unfounded: Research on safety in the shipping industry by Baniela and Rios (2010, 2013) showed evidence that risk homeostasis could be involved in accident rates that did not seem to drop despite technological advancements and implemented safety measures. Further inspection led the way back to important factors in Risk Homeostasis Theory: costs and benefits. Partly 'on the floor' but especially on management level, the (mostly financial) benefits of risky choices outweighed their costs, leaving optimal safety as the less profitable option. This process of end-over-means is a generally occurring element in risk-taking, as earlier described by Apter (1984). Despite these findings, Risk Homeostasis has yet to enter the occupational field as a mechanism considered in safety measures (Swuste et al., 2017).

Risk homeostasis has been found in digital behavior contexts as well. A study by Sawyer et al. (1999) showed that the threat of a computer virus lead users to take more protective measures than they did before. The proximity of the threat (personal experience, learning through the media) moderated the intensity of this reaction.

Aside from situations with a risk of immediate (physical) harm, RHT was also involved in studies on health decisions (Maughan-Brown & Venkataramani, 2012) and consumer behavior (Miller, 1998).

1.3.2 Research critical of RHT

As discussed earlier, criticism on the Risk Homeostasis Theory is formulated on several grounds, with its methodological shortcomings as the main issue. In 1986, Evans investigated risk homeostasis by studying accident rates around the implementation of a law enforcing the use of motorcycle helmets. In addition to finding contradictory evidence, Evans expressed having issues with the lack of a concrete way to measure the homeostasis process. Without a clear definition of homeostasis factors, changes in accident rates could just as well be influenced by ‘selective recruitment’ (Evans, 1986): drivers who actually comply to the law can be expected to behave more safely than their bare-headed counterparts by default. McKenna (1987) shared this view. In his rebuttal of RHT research, he mentioned the issue of unmeasurable objective risk, as well as how the theory cannot control for all factors realistically involved in accidents.

Janssen and Tenkink (1988), and later Trimpop (1996) propose the theory be refined by expanding the number of factors involved and developing a way to measure actual homeostasis, after which it could be properly confirmed or falsified. After all, the lack of a concrete and unambiguous process description would lead to a theory that is confirmed under all circumstances, and thus unfalsifiable (Adams, 1988; Hoyes & Glendon, 1993; Elvik, 2004). Without clearly defined conditions for falsification, research against RHT consists mainly of a lack of support.

Like Evans in 1986, Shannen and Szatmari (1994) researched accident rates, this time during the introduction of a law requiring drivers to wear seatbelts. Following Risk Homeostasis Theory, the expectation would be for people to start driving more recklessly to compensate for their newly-gained safety. This did not seem to occur as no significant difference in accidents was found before and after the seatbelt regulations. The earlier mentioned driving simulator study by Hoyes et al. (1996) produced contradictory results as it was specifically aimed at the role of utility in risk homeostasis. Although the RHT describes utility as a catalyst for risky behavior, no significant homeostasis took place in gain-related risk-taking behavior. O’Neill (1998) enumerates RHT research with disconfirming results in a review refuting the theory and its followers. He mentions research by Haight (1986) and Evans (1986) containing the conclusion that risk homeostasis as proposed by Wilde (1982) is little more than a ‘philosophical claim’ (Evans, 1991).

In light of these arguments, many opportunities exist to further refine the Risk Homeostasis Theory and its research methods to move towards being a concrete, measurable entity that is falsifiable enough for most to agree on.

1.4 Risk-taking, Music and Emotion

A field of study relevant to the present study is music preference and its relation to risk-taking. Music is an inseparable part of human life, used in various situations throughout our daily activities. Although it has been shown that practicing music is related to our cognitive toolset and personality (Corrigall et al., 2013), experiencing it from a listening perspective could carry different implications. Since people are said to consume around 4 hours of audio a day (Webster, 2014), this perspective is of considerable size. According to a study by North et al. (2004), people listen to different kinds of music, depending on their environment. In this study, people reported listening to music they preferred in private situations, for example when driving. When in larger groups, music played was liked generally and didn't necessarily correspond to the preferences of the group members.

1.4.1 Music preference

The relationship between a person's music preference and their risk-taking behavior has been addressed in research previously. Gregersen and Berg (1994) performed a questionnaire study on young Swedish drivers and their lifestyle. Activities, preferences and motivations were associated with several high and low risk groups. In this study, activities related to culture and music explained a medium amount of risk-taking, but was not specifically distinct in a risk group. Schulze (1990, as discussed in Gregersen & Berg, 1994) had performed a similar study, dividing German adolescents into groups according to their accident involvement. One prominent high-risk group was characterized by a preference for rock and punk music, time spent away from home and alcohol consumption. Arnett (1992) compared adolescents with a preference for heavy metal to those without. This comparison showed that those who liked heavy metal music behaved more recklessly on several aspects such as vandalism, sexual relations, drug use and driving. These adolescents were also more sensation-seeking. In a recent study by Enstrom and Schmaltz (2017), music preference was linked to different kinds of risk-taking. According to the results, those who preferred music with an 'aggressive' sound were more eager to take action-oriented, recreational risks such as bungee jumping. A preference for music with a 'mellow' sound was correlated with more social risk-taking, for instance offering a bold opinion or moving far away from family.

1.4.2 Experiencing preferred music and risk-taking

Associations between different kinds of preferred music and risk-taking have been found, but are mixed. When talking about preference as a trait, the question is: how does experiencing this preference influence a person's risk-taking behavior? This was addressed in studies by Halko et al. (2015) and Halko and Kaustia (2015). In these studies, young adults were asked to perform sets of gambles with real money. Before the experiment, they picked out different musical pieces as either liked or disliked. These pieces were then played to participants during the experiment. The results showed that participants took more risks when listening to liked music, and less when listening to disliked music. A possible explanation proposed by the writers was that listening to preferred music induced a positive mood, which in turn influenced risk-taking behavior. Halko and Kaustia (2015) described that a positive mood might influence people's outlook on risk-taking, leading to underestimation of the consequences. Additionally, an explanation in terms of Kahneman and Tversky's (1979) Prospect Theory was given: listening pleasure increases experienced utility, making future gains seem smaller as a part of current gains and reducing risk aversion.

1.4.3 Musical aspects and risk-taking

Aside from personal preferences and the effect of experiencing these, smaller attributes of music have been linked to risk-taking as well. Ayres & Hughes (1986) found that loud background music (107 dB) reduced drivers' visual acuity. In a driving simulator experiment by Brodsky (2002), participants had a significantly higher speed and more violations when listening to fast music. Interestingly, results from a study by Spinney (1997) indicated that the presence of background music improved drivers' reaction time and decreased the number of accidents. This suggests that the relation between music and risk-taking might be of non-linear and complex nature.

1.4.4 Musical influence through emotions

In the studies by Halko et al. (2015) and Halko and Kaustia (2015), the influence of music on risk-taking seemed to (partially) occur through an emotional channel. Kreutz et al. (2008) used 25 pieces of classical music, each associated with a different emotion, to induce these emotions in participants. The results showed that it was possible to arouse emotions in humans with only exposure to music, although this effect was strongest for positive emotions such as 'happy' and 'peaceful'. According to a model by Juslin (2013), 'aesthetic' emotions present in music are converted to 'real-life' emotions we are capable of feeling through signals of meaning the brain tries to derive from the sound, as if it were a language. However, there is some nuance to this analogy. As Juslin (2013) and North et al. (2004) noted, different people feel and interpret music in different ways: the exact same sound can be received differently depending on the ears receiving them.

The effect of these emotions on risk-taking was assessed by Yuen and Lee (2003). Participants were shown a video clip with either sad, neutral or happy contents. After that, they were asked to fill out a questionnaire on risky decisions. Participants in a sad mood took significantly less risks than the other mood groups, with an increase in risk as moods became happier. Stalder and Cook (2014) added to these findings in a revision of Forgas' (1998) theory on the Fundamental Attribution Error. When in a happy mood, people tended to judge situations and other people wrongly. According to Forgas (1998), this is caused by a lack of attention to detail when happy. These errors in judgment can lead to bad decisions, increasing risk.

The extent to which emotions are absorbed depends on an individual's susceptibility to them. Hatfield et al. (1993) refer to this as 'emotional contagion'. The higher the emotional contagion, the quicker a person 'catches' emotional states from their environment. This process relies on the premise that the emotion is noticeable emotion and can be responded to (Dezecache et al., 2015). In the context of emotions as a mediator between music and risk-taking, emotional susceptibility represents the factor of intensity. In that case, a highly susceptible person would be expected to have a larger amount of emotional influence mixed into their decision-making, and thus their risk-taking.

1.5 Present study

This study will be performed to learn more about the workings of the Risk Homeostasis Theory, and add clarity to the controversial body of existing research around it. To this end, an experiment will be carried out in controlled settings, taking into account the theory's measurement and falsifiability issues. This will be in the shape of a computer game in which flying a rocket through a field of meteors to gain points is the player's goal. In this game, protection is represented by the amount of shields the rocket has to endure a collision with a meteor. Risk-taking is expressed in flying speed and the time and distance kept to nearby meteors.

By using concrete definitions of the independent factor 'protection' and dependent factor 'risk-taking', the relation between the two and the occurrence of a homeostatic effect in a game environment can be investigated closely. These results can then be used to carefully seek insight into real-life risk homeostasis and the effectiveness of safety measures.

Following Wilde's (1982) Risk Homeostasis Theory, expectations are that risky behavior will increase as the amount of protection goes up. In the context of risk homeostasis, this would mean that higher levels of protection lower an individual's amount of perceived risk while their desired risk remains stable, creating a gap. To close the distance, compensatory risk would then be taken to return to a state of risk equilibrium.

Additionally, the role of music in risk-taking behavior will be considered. This will be analyzed in the shape of genre preference and exposure in hours per week. It is expected that different preferences in music genre are connected to different levels of risk-taking. More hours of weekly exposure are expected to lead to more risk-taking, as well. For music preference, research on its relation to risk-taking has been scarce and with mixed outcome (Schulze, 1990; Arnett, 1992; Enstrom & Schmaltz, 2017). The results of this study will add to existing findings and help create a clearer picture of musical risk-taking. The role of exposure could add to this picture as an intensity aspect. A larger understanding of music as a factor in risk-taking can better be taken into account for possible safety measures.

Since music has been linked to emotions in several ways (Halko et al., 2015; Kreutz et al., 2008; Juslin, 2013), emotional susceptibility will be used as a nuancing factor in the music-risk relation. Higher susceptibility to emotion is expected to relate to higher levels of risk-taking, as musical influences on emotion would hit susceptible people harder. Knowledge on emotion as a mediator of music as well as other factors to behavior is of great value in enhancing rules, training and prevention around risks and accidents.

This study aims to answer two main questions:

1. How does the amount of protection influence risk-taking and is a homeostatic effect involved?
2. What is the relation between musical preference and exposure to risk-taking, and how does emotional susceptibility mediate this?

To answer these questions, the following hypotheses were formulated:

Hypothesis 1a: A higher level of noticeable protection is accompanied by a higher measure of risk-taking behavior.

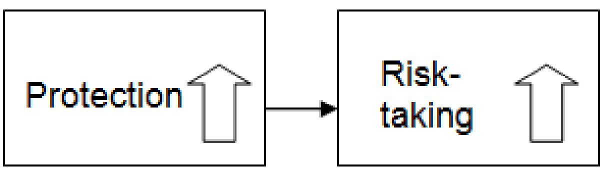

Figure 2. Schematic representation of hypothesis 1a.

Hypothesis 1b: There is a difference in effectiveness of risk-taking strategies.

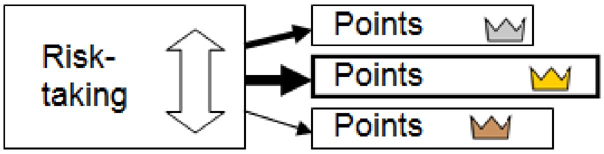

Figure 3. Schematic representation of hypothesis 1b.

Hypothesis 2a: More hours of listening to music causes more risk-taking behavior.

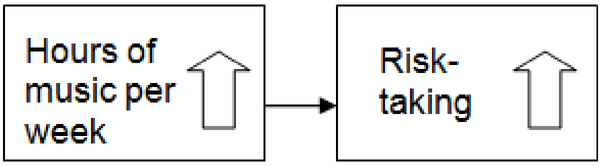

Figure 4. Schematic representation of hypothesis 2a.

Hypothesis 2b: Emotional susceptibility has a mediating effect on the relation between hours of listening to music and risk-taking behavior.

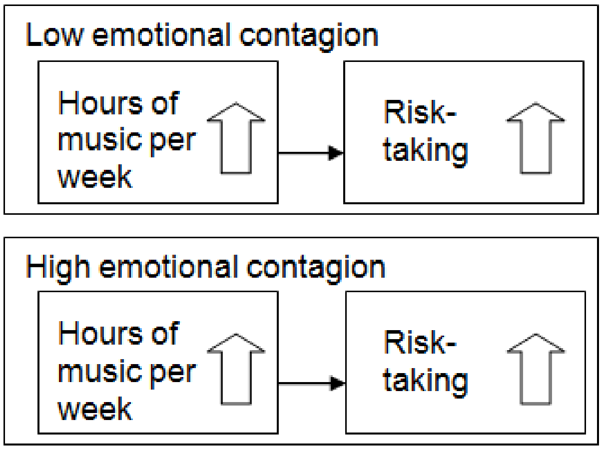

Figure 5. Schematic representation of hypothesis 2b.

Hypothesis 3a: Different music preferences lead to different risk homeostasis effects / different amounts of risk-taking behavior

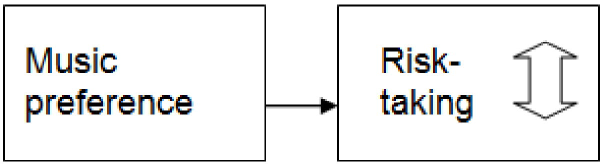

Figure 6. Schematic representation of hypothesis 3a.

Hypothesis 3b: Emotional susceptibility has a mediating effect on the relation between music preference and risk-taking behavior.

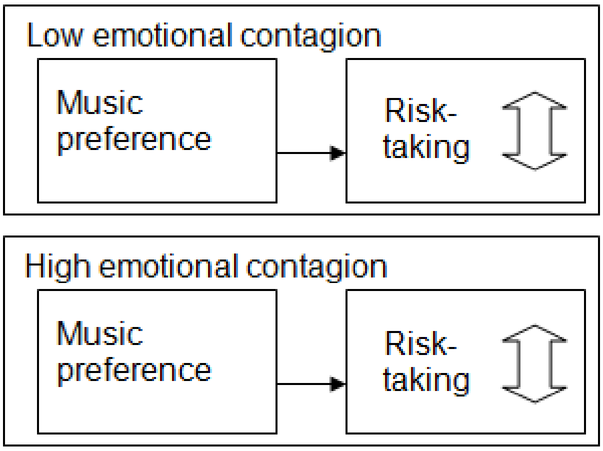

Figure 7. Schematic representation of hypothesis 3b.

The results of this study hold value as an addition to the current body of research on the Risk Homeostasis Theory. Due to its controversial nature, findings up until now are all but uniform. Hopefully, controlling for as many factors as possible in this laboratory experiment will produce valuable insights on risk homeostasis and contribute to proper validation. Aside from findings on risk homeostasis, results of the music and emotional susceptibility analyses might help expand the realm of knowledge around music as an indicator of risk-taking behavior and its mechanisms. Additionally, new findings could lead to a better understanding of current safety precautions and their effectiveness. This understanding can help shape future measures even more accurately.

2. Methods

2.1 Participants

The 69 participants in this study were between 18 and 36 years of age (M = 22,4, σ = 3,22). 58 were female (84,1%), 11 were male (15,9%). Highest completed education ranged from HAVO (high school; 'Hoger Algemeen Voortgezet Onderwijs') to WO Master (University master's degree). VWO (high school; scientific track) (N=33) and WO Bachelor (university bachelor's degree) (N=23) were the most represented.

Participants were recruited through multiple channels. The main means of recruiting was Leiden University's student research participation system, SONA. Through SONA, 62 students signed up for the research. After cancellations and no-shows, 51 of these participants completed the experiment. Additional participants were recruited through the researchers' social circles.

Exclusion criteria for this study were the presence of one or more psychological and/or neurological disorders and previous experience with the game used in the experiment.

No technical or data processing problems arose during the study. The sample contained no missing data.

2.2 Materials

Two means for gathering data were used: a questionnaire and a video game.

The questionnaire was used for multiple fields of research, and contained questions on subjects such as diet, perceived masculinity, substance use and sports. Relevant to this study, questions were included for collecting data on participants' music preference and emotional susceptibility. Questions on music preference were self-developed, as no questionnaire or general way of measure existed at the time (appendix A). Participants were asked for three favorite music genres from a list, and could enter unrepresented genres manually through an 'other'-option. Emotional susceptibility was measured through the Emotional Contagion Scale, developed by Doherty (1997) (appendix B).

The questionnaire was built and filled in using Qualtrics. The questions were in English.

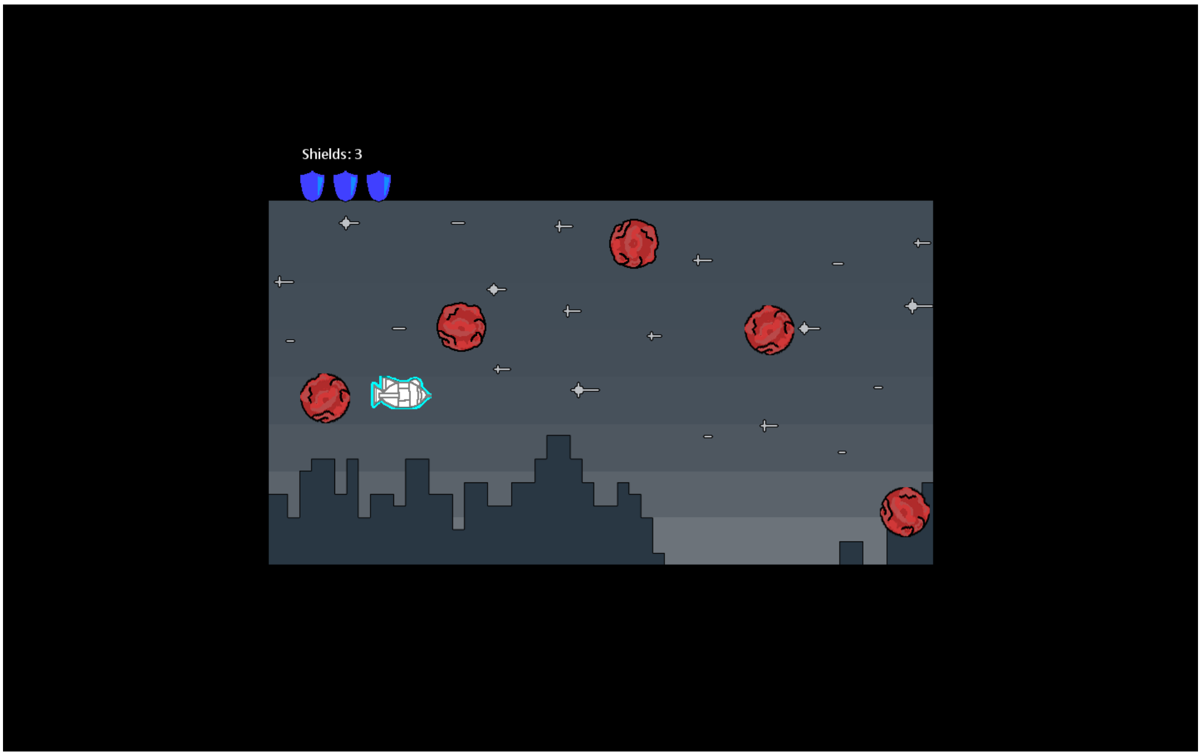

Figure 8. A screenshot of The Spaceship Game.

The video game

The game was used for measuring risk behavior in a virtual setting; The Spaceship Game. In the game, the player's goal was to 'deliver important cargo' in a spaceship while avoiding meteors. This meant navigating the spaceship through a field of infinite meteors appearing from the right of the screen, as illustrated in figure 8.

Navigation happened through the keyboard's arrow buttons; the 'up' and 'down' buttons moved the spaceship vertically, while 'left' and 'right' were used to control flying speed. Speed could be varied on 13 levels, starting at 320 pixels per second with a maximum of 920 pixels per second, increasing by 50 pixels per difficulty level. Each time the player hit a meteor, one shield was subtracted. Hitting a meteor with zero shields ended the session. The amount of shields was visible to the player in the top-left corner.

Game environment

At the start of the game, the player was shown a preview video of the gameplay. A test round was played to allow the player to become accustomed to the navigation. The practice round knew two variations: one with zero shields and one with three shields. After the test round, five sessions were played. These were the 'test conditions'. During these sessions, the player started out with zero, one, three, four or five shields. The order of the shield amount was randomly assigned from 120 different orders. Each player was assigned a unique session order. A session lasted for four minutes, or until the player ran out of shields.

Data

The game data was recorded in the files 'steplog.csv' and 'eventlog.csv'. The first contains data for each 0.1 second, while the latter contains data for the moments the spaceship collided and a session ended. The registered parameters were: participant number, session number, time elapsed since the start of the session, player score in points, speed level, amount of shields, horizontal position of the spaceship from the left border, vertical position of the spaceship from the top border, the meteor closest on the x-axis and y-axis and *whether a collision was taking place.*

Three risk-taking parameters were calculated using data from 'steplog.csv': speed, distance kept to the closest meteor (DCM) in pixels and the time to collision with the closest meteor in its path (TTC) in seconds, if the ship would fly straightly forward. These are illustrated in figure 9. The number of shields was used as a parameter for the player's perceived amount of protection. In relation to each other, a high speed and low DCM and TTC indicate a high amount of risk-taking. A low speed and high DCM and TTC indicate a low amount of risk-taking.

Specific details on variable calculation can be found in appendix C.

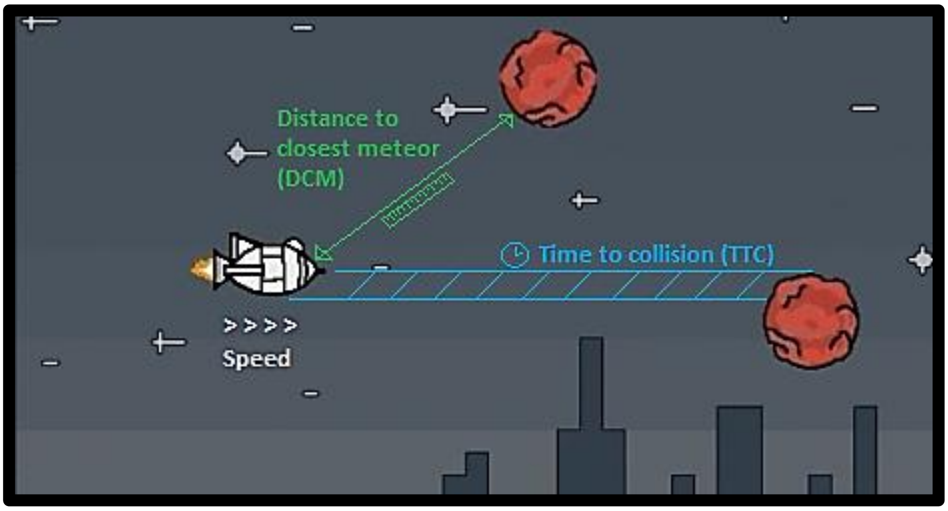

Figure 9. Visual representation of the risk-taking parameters.

2.3 Design

For this study a double-blind randomized design was used. Five conditions started with 0, 1, 3, 4 or 5 shields were randomly presented to each participant. The order of these conditions was unknown to both participant and researchers.

2.4 Procedure

Environment

The tests were performed in the computer room of Leiden University's Faculty of Social Sciences (FSW). Testing lasted four days, during which at least two researchers were present. Completing the questionnaire and game took participants around 45 minutes. The computers were identical in hardware specifications and screen size **(**21"**)**. The questionnaire and the game both contained no sound.

Procedure

At the start, participants were given an instructions sheet and a consent form to sign (see appendices D and E). These papers contained general information about the test procedure and the reassurance that data was coded anonymously and that participants could stop anytime during the test. After signing the consent form, the game was started by one of the researchers. Instructions for playing, as well as a preview video, were included in the game and shown at the beginning. While playing the game participants could acquire points. These were invisible to the player. The amount of points gained was determined by the difficulty level(speed) used, and the total time played. The three participants in first, second and third place were awarded an extra prize of €50, €20 and €10 respectively. These participants were contacted approximately a week after the last test. After finishing the game, participants could raise their hand to have the questionnaire put on. When the game and questionnaire had been completed, participants were awarded €6,50 or 2 student participation credits.

2.5 Analyses

Variables

Analyses were conducted with IBM SPSS Statistics 23. The data was checked for missing cases, but none existed. All data was analyzed. Questionnaire data from Qualtrics was recoded into variables. Data from the game's 'steplog' file was recoded for analysis as well.

Variables calculated from Qualtrics were:

1. Favorite music genre, divided into categories 'pop', 'rock', 'electronic' and 'other'
2. Hours of listening to music per week, divided into categories of '0 - 0,5 hours', '0,5 - 2 hours', '2 - 4 hours', '4 – 6 hours' and 'more than 6 hours'.
3. Emotional contagion score, a numerical score from 15 to 60.

Because 'hours of music' group '0 – 0,5 hours' contained no cases, it was removed.

Since no official division of music genres has been recorded, the categories in the questionnaire were established by consulting several websites ("The fundamental music genre list", 2011; "Music Genre list", 2015) and the Discogs database ("Discogs – Database and…", 2016), since this is an internationally widely-used medium. A selection was made from the genres that appeared most regularly. To keep the answer options as complete as possible, an 'other' option was provided with a field to enter a genre manually. These entries were recoded into existing categories, or a new category was created. This lead to a wide amount of reported favorites. To be able to analyze the music preference groups, those with a substantial size were used in the analyses, combining the groups too small to analyze into the 'other' category.

For risk-taking behavior, speed, time to collision (TTC) and distance to the closest meteor (DCM) were calculated from the step log file and used as variables (appendix C).

Hypothesis 1a was tested by conducting a repeated measures ANOVA. By conducting an RMA instead of multiple paired-samples t-tests, the probability of a Type I-error is lowered, thus making this option more reliable. The means of risk-taking parameters speed, DCM and TTC were analyzed for significant differences. These differences were tested between conditions for a long-term effect and within conditions for a short-term effect. The assumptions for a repeated measures ANOVA were checked. These are discussed further in 'Results.'

Hypothesis 1b was to be tested using Pearson's correlation. Because the relation between variables turned out to be non-linear, Spearman's test and Kendall's tau-b were used instead.

Hypotheses 2a and 3a were tested by performing one-way ANOVAs. A MANOVA was originally intended, but the dependent variables were too closely correlated, as well as not linearly related.

Hypothesis 2b and 3b were planned to be tested by performing one-way ANCOVAs. However, checking for assumptions of this test raised many issues, suggesting that emotional susceptibility is not suited as a covariate of risk-taking. Therefore, several repeated-measures ANOVAs were conducted, splitting the file into 'low susceptibility' and 'high susceptibility'. These groups' risk behavior was compared. For the one-way ANOVAs as well as the repeated measures ANOVAs, assumptions were checked. Assumptions that could only be checked during the analysis will be reported alongside the results.

3. Results

3.1 Hypothesis 1: Risk homeostasis and strategy

3.1.1 Assumptions tests

For the repeated measures ANOVA, assumptions were checked. Since the dependent variables are all measured on a continuous scale, assumption 1 is met. Assumption 2 is met since the within-subjects factor is categorical with two or more levels. Boxplots were used to find significant outliers in any level of the within-subjects factor. These outliers were found, but not removed since they are relevant to the researched phenomenon. Normal distribution of dependent variables for each level of the within-subjects factor was tested with Q-Q plots. The Shapiro Wilks-test was not used for this assumption, because the sample size was larger than 50, reducing its reliability. Normal distribution in all of the dependent variables. The assumption of sphericity was assessed with Mauchly's test during the analyses. Its results and possible corrections will be mentioned along with each result.

For the Pearson's correlation in hypothesis 1b, the assumptions of the variables being continuous and paired were met. The linearity assumption was violated, according to visual inspection of scatterplots (figure 10). Transformation of the variables still did not result in linearity of the variables. Therefore, Pearson's correlation could not be used and assumptions for a non-parametric Spearman's test were checked.

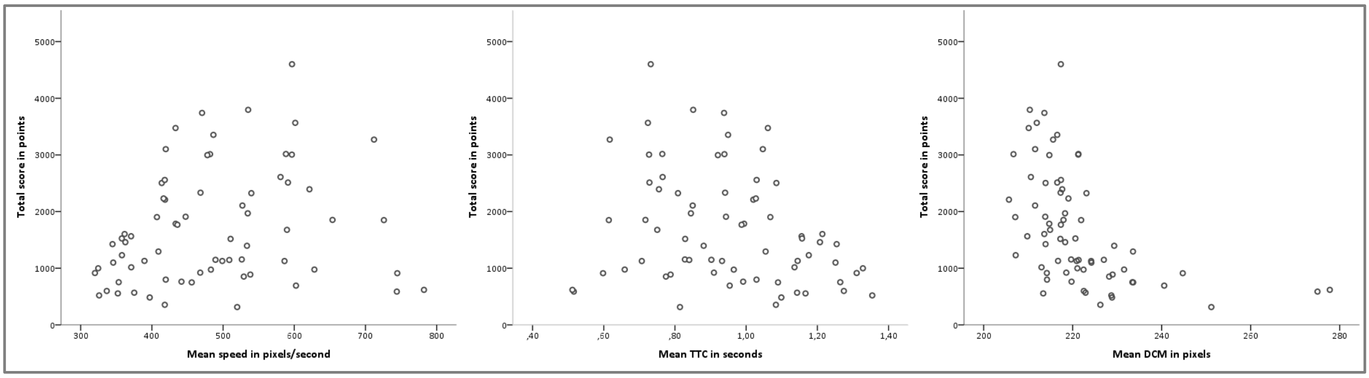

Figure 10. Scatterplot mean risk parameters x total score.

The assumptions for continuous and paired variables were met. The assumption of a monotonic relationship between the variable pairs was met partly: only the 'Mean DCM – Total score' pair contained a monotonic relationship (figure 10). This pair was included in the Spearman's correlation.

For the ‘Mean speed – Total score’ and ‘Mean TTC – Total score’ pairs, a Kendall’s tau-b association test was performed. The assumptions of continuity and pairing of the variables were met, as in the Pearson and Spearman assumptions checks. The third assumption, the appearance of direction, was accepted as the data in the scatterplots formed a pattern to be further analyzed for association.

3.1.2 Hypothesis 1a: Risk homeostasis between conditions

One-way repeated measures ANOVAs were used to analyze players' speed, DCM and TTC over the five conditions (started with 0 shields, 1 shield, 3 shields, 4 shields and 5 shields).

Speed

Descriptive statistics for each of the 5 conditions can be found in table 1.

	M	SD	N
0-shield condition	432,49	112,42	69
1-shield condition	467,36	121,77	69
3-shield condition	503,39	139,6	69
4-shield condition	521,22	160,73	69
5-shield condition	528,61	142,35	69

Table 1. Descriptive statistics between conditions: Speed.

The assumption of sphericity was met, $X^2(9) = 11.508$, $p = .243$. The multivariate tests were significant ($p < .001$).

The univariate result showed a significant difference between the speed means of the 5 conditions, $F(4,272) = 17.44$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there were significant differences between the conditions with 0 and 3 shields (p = < .001), 0 and 4 shields (p = < .001), 0 and 5 shields (p = < .001), 1 and 4 shields (p = < .001), and 1 and 5 shields (p = < .001). This is represented visually in table 2.

The linear relation of speed and amount of shields at the start is shown in figure 11. The trend suggests that players' mean speed increases as the session started with more shields.

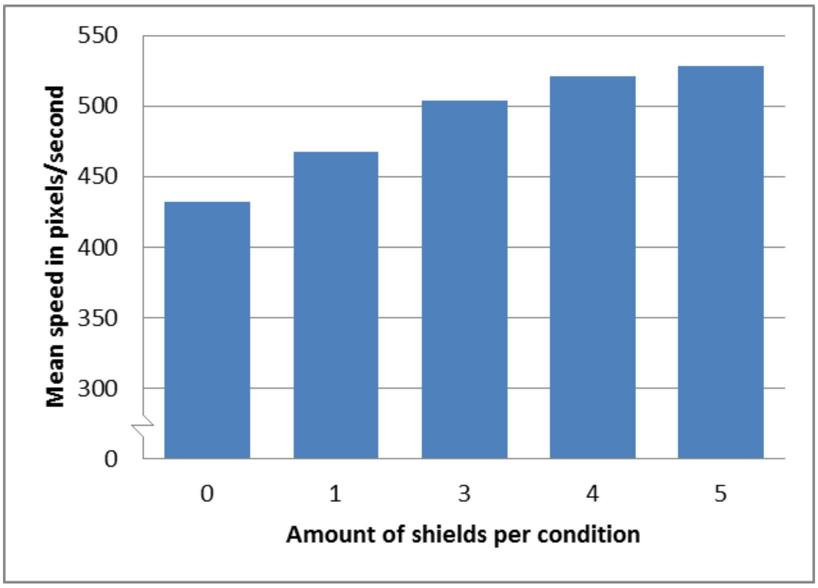

Figure 11. Mean speed between conditions.

Shields	0	1	3	4	5
0		-	*	*	*
1			-	*	*
3				-	-
4					-
5					

Table 2. Significant differences between conditions: Speed.

TTC

The assumption of sphericity was violated, $X^2(9) = 18.855$, $p = .026$. Thus, degrees of freedom for the univariate results were corrected with the Huynh-Feldt estimates of sphericity ($\varepsilon = .942$).

The multivariate tests were significant ($p < .001$). The univariate result showed a significant difference between the TTC means of the 5 conditions, $F(3.766, 252.349) = 17.415$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there were significant differences between the conditions with 0 shields and 1 shield ($p = .026$), 0 and 3 shields ($p < .001$), 0 and 4 shields ($p < .001$), 0 and 5 shields ($p < .001$), 1 and 4 shields ($p = .012$) and 1 and 5 shields ($p < .001$).

The linear relation of speed and amount of shields at the start is shown in figure 12. The trend suggests that players' mean TTC decreases as the session started with more shields.

	M	SD	N
0-shield condition	1,06	0,27	68
1-shield condition	0,97	0,23	68
3-shield condition	0,91	0,25	68
4-shield condition	0,88	0,27	68
5-shield condition	0,85	0,23	68

Table 3. Descriptive statistics between conditions: TTC.

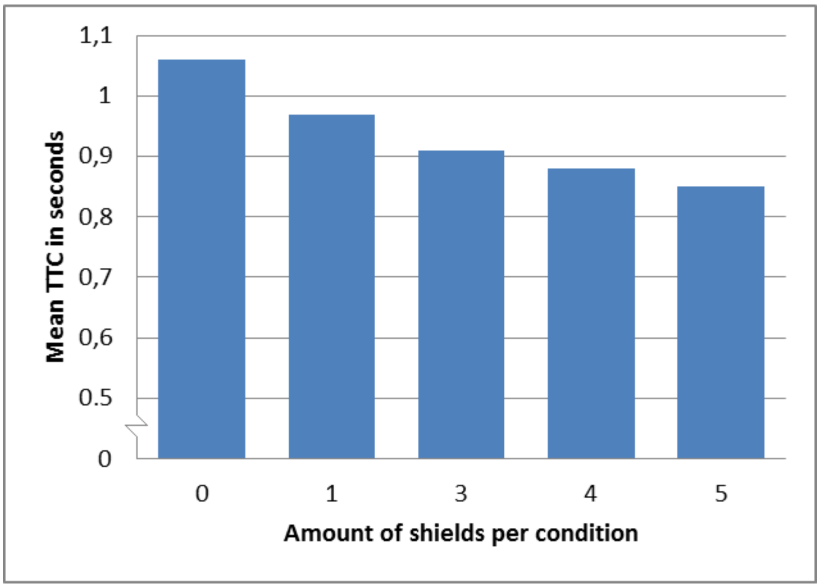

Figure 12. Mean TTC between conditions.

Shields	0	1	3	4	5
0		*	*	*	*
1			-	*	*
3				-	-
4					-
5					

Table 4. Significant differences between conditions: TTC.

DCM

Mauchly's test showed that the assumption of sphericity was violated, $X^2(9) = 321.123$, $p < .001$. Thus, degrees of freedom for the univariate results were corrected with the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .336$).

The multivariate tests were significant ($p < .001$). The univariate result showed a significant difference between the TTC means of the 5 conditions, $F(1.345, 91.462) = 21.662$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there were significant differences between the conditions with 0 shields and 1 shield ($p = .002$), 0 and 3 shields ($p < .001$), 0 and 4 shields ($p < .001$), 0 and 5 shields ($p < .001$), 1 and 4 shields ($p = .011$), 1 and 5 shields ($p = .001$) and 3 and 5 shields ($p = .023$).

The linear relation of DCM and amount of shields at the start is shown in figure 13. The trend suggests that players' mean DCM decreases as the session started with more shields.

	M	SD	N
0-shield condition	261,89	71,55	69
1-shield condition	231,33	27,65	69
3-shield condition	223,97	15,95	69
4-shield condition	221,08	12,56	69
5-shield condition	218,80	11,48	69

Table 5. Descriptive statistics between conditions: DCM.

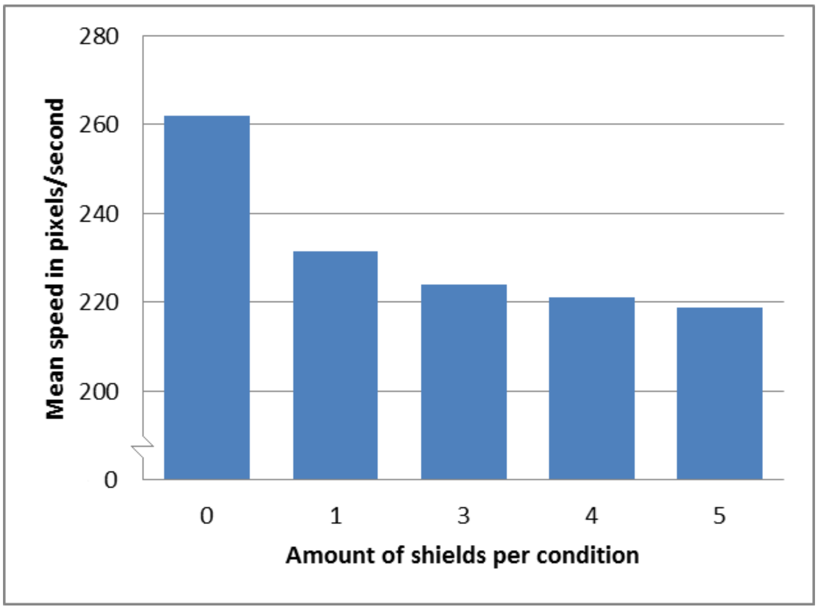

Figure 13. Mean DCM between conditions.

Shields	0	1	3	4	5
0		*	*	*	*
1			-	*	*
3				-	-
4					-
5					

Table 6. Significant differences between conditions: DCM.

3.1.3 Hypothesis 1a: Risk homeostasis within conditions

To test the occurrence of risk homeostasis on a shorter term, repeated measure ANOVAs were also conducted within the conditions. Speed, TTC and DCM were analyzed for each shield that was left during a condition to see if players' behavior had become more cautious for each shield they lost.

TTC and DCM are not included in all analyses. These parameters were calculated using the in-game location of meteors, which don't appear until a short time after the condition starts, causing a small data distortion during the first shield. For the one-shield condition, containing only 2 shield options (1 or 0), this data was considered to be too clouded to be reliable. For the other conditions, the first shield was left out of the TTC and DCM analyses.

3.1.3.1 Within conditions - One shield

For the one shield-condition, only speed has been analyzed.

Speed :

Testing for sphericity is not applicable since only two conditions were compared. The multivariate tests were significant (p = .001).

The univariate result showed a significant difference between the two conditions, $F(1, 68)$ = 51.994, $p < .001$. Speed is higher with no shields left than with one shield left.

	N	M	SD
1/1 shields	69	440,20	110,43
0/1 shields	69	524,69	164,87

Table 7. Descriptive statistics within 1-shield condition: Speed

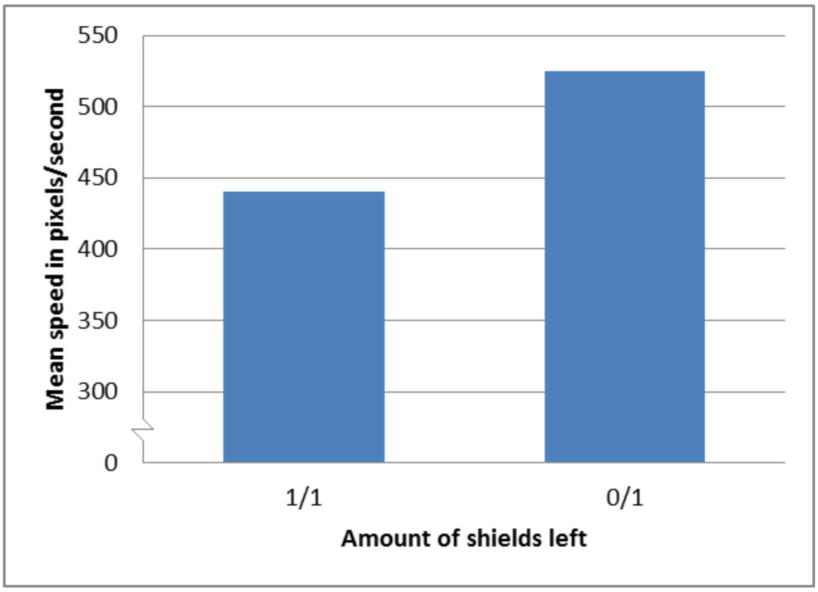

Figure 14. Mean speed within conditions: 1-shield condition.

3.1.3.2 Within conditions - Three shields

Speed:

The assumption of sphericity was violated, $X^2(5) = 62.85$, $p < .001$. Thus, degrees of freedom for the univariate results were corrected with the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .611$).

The multivariate tests were significant ($p < .001$). The univariate result showed a significant difference in mean speed between the shields left within conditions, $F(1.833, 104.461) = 49.576$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there was significant difference in the mean speed between 3 and 2 shields left ($p < .001$), 3 shields and 1 shield left ($p < .001$), 3 and 0 shields left ($p < .001$), 2 shields and 1 shield left ($p = .001$) and 2 and 0 shields left ($p = .001$). This is illustrated in table 9.

	N	M	SD
3/3 shields	58	458,06	113,27
2/3 shields	58	556,83	171,11
1/3 shields	58	596,79	185,98
0 shields	58	608,20	179,44

Table 8. Descriptive statistics within 3-shields condition : Speed.

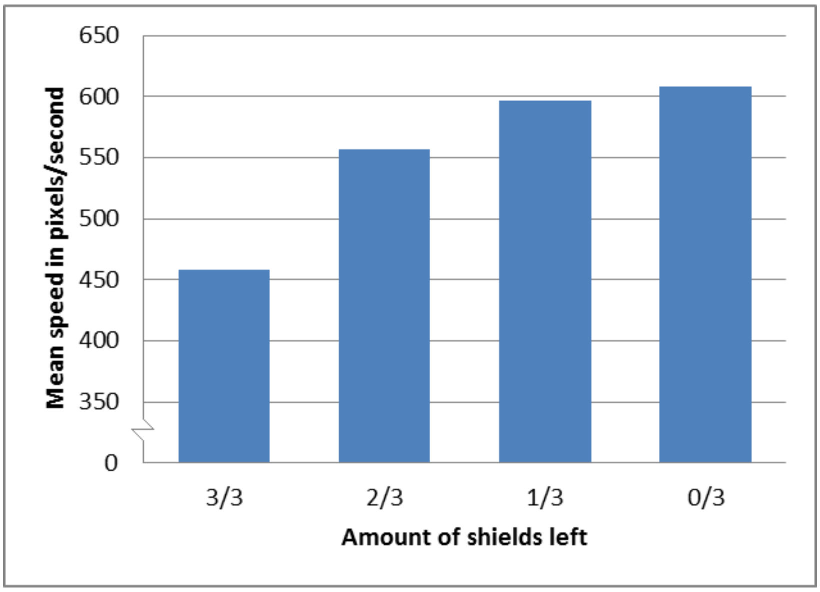

Figure 15. Mean speed within conditions: 3-shield condition.

Shields left	3/3	2/3	1/3	0/3
3/3		*	*	*
2/3			*	*
1/3				-
0/3				

Table 9. Significant differences in speed within conditions: 3-shield condition.

TTC:

The assumption of sphericity was met, $X^2(2) = 2.211$, $p = .900$.

The multivariate tests were not significant ($p = .126$). The univariate result showed no significant difference in mean TTC between the shields left within conditions, $F(2, 114) = 2.239$, $p = .111$. Thus, no post-hoc tests were conducted.

	N	M	SD
2/3 shields	58	0,75	0,34
1/3 shields	58	0,68	0,31
0/3 shields	58	0,69	0,29

Table 10. Descriptive statistics within 3-shields condition: TTC.

DCM:

The assumption of sphericity was met, $X^2(2) = 2.874$, $p = .238$.

The multivariate tests were not significant ($p = .640$). The univariate result showed no significant difference in mean TTC between the shields left within conditions, $F(2, 114) = 0.362$, $p = .697$. Thus, no post-hoc tests were conducted.

	N	M	SD
2/3 shields	58	203,31	16,43
1/3 shields	58	202,83	16,21
0 shields	58	205,08	12,37

Table 11. Descriptive statistics within 3-shields condition: DCM.

3.1.3.3 Within conditions - Four shields

Speed:

The assumption of sphericity was violated, $X^2(9) = 113.003$, $p < .001$. Thus, degrees of freedom for the univariate results were corrected with the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .515$).

The multivariate tests were significant ($p < .001$). The univariate result showed a significant difference in mean speed between the shields left within conditions, $F(2.059, 115.313) = 36.817$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there was significant difference in the mean speed between 4 and 3 shields left ($p < .001$), 4 and 2 shields left ($p < .001$), 4 and 1 shields left ($p < .001$) and 4 and 0 shields left ($p < .001$). This is illustrated in table 13.

	N	M	SD
4/4 shields	57	473,79	142,83
3/4 shields	57	578,07	195,87
2/4 shields	57	586,18	178,60
1/4 shields	57	594,91	184,60
0/4 shields	57	595,49	183,13

Table 12. Descriptive statistics within 4-shields condition: Speed.

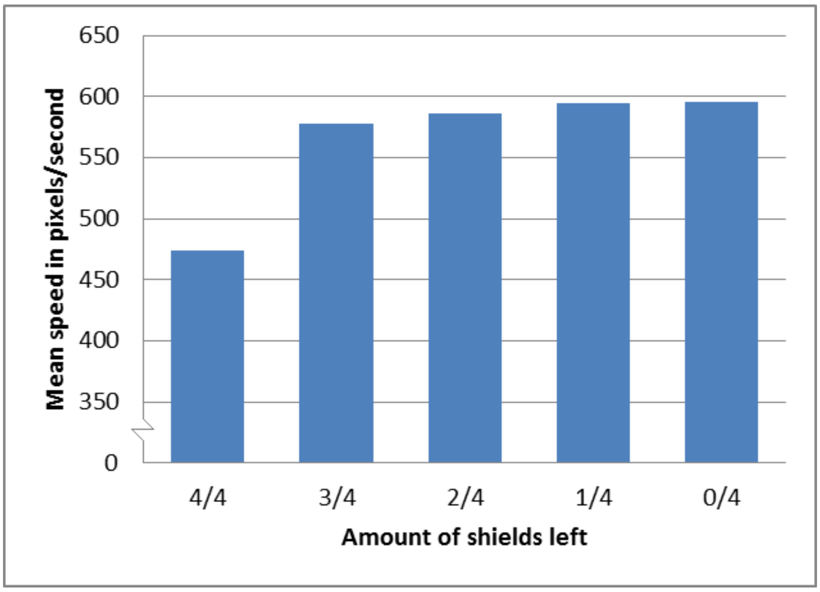

Figure 16. Mean speed within conditions: 4-shield condition.

Shields left	4/4	3/4	2/4	1/4	0/4
4/4		*	*	*	*
3/4			-	-	-
2/4				-	-
1/4					-
0/4					

Table 13. Significant differences in speed within conditions: 4-shield condition.

TTC:

The assumption of sphericity was violated, $X^2(5) = 16.530$, $p = .005$. Thus, degrees of freedom for the univariate results were corrected with the Huynh-Feldt estimates of sphericity ($\varepsilon = .881$).

The multivariate tests were not significant ($p = .816$). The univariate result showed no significant difference in mean TTC between the shields left within conditions, $F(2.642, 147.965) = 0.231$, $p = .852$. Thus, no post-hoc tests were conducted.

	N	M	SD
3/4 shields	57	0,74	0,28
2/4 shields	57	0,73	0,29
1/4 shields	57	0,71	0,27
0/4 shields	57	0,73	0,27

Table 14. Descriptive statistics within 4-shields condition: TTC.

DCM:

The assumption of sphericity was met, $X^2(5) = 4.149$, $p = .528$.

Multivariate tests were not significant ($p = .825$). The univariate result showed no significant difference in mean DCM between the shields left within conditions, $F(3, 171) = 0.263$, $p = .852$. Thus, no post-hoc tests were conducted.

	N	M	SD
3/4 shields	58	207,95	17,37
2/4 shields	58	206,97	14,49
1/4 shields	58	207,33	17,08
0/4 shields	58	209,24	14,52

Table 15. Descriptive statistics within 4-shields condition: DCM.

3.1.3.4. Within conditions – Five shields

Speed:

The assumption of sphericity was violated, $X^2(14) = 214.663$, $p < .001$. Thus, degrees of freedom for the univariate results were corrected with the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .501$).

The multivariate tests were significant ($p < .001$), and the univariate result showed a significant difference in mean speed between the shields left within conditions, $F(2.505, 155.285) = 36.087$, $p < .001$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there was significant difference in the mean speed between 5 and 4 shields left ($p < .001$), 5 and 3 shields left ($p < .001$), 5 and 2 shields left ($p < .001$), 5 and 1 shields left ($p < .001$), 5 and 0 shields left ($p < .001$) and 4 and 2 shields left ($p = .001$). This is illustrated in table 17.

	N	M	SD
5/5 shields	63	456,74	119,29
4/5 shields	63	555,47	172,89
3/5 shields	63	576,11	183,92
2/5 shields	63	585,34	181,86
1/5 shields	63	590,94	172,34
0/5 shields	63	592,98	184,65

Table 16. Descriptive statistics within 5-shields condition: Speed.

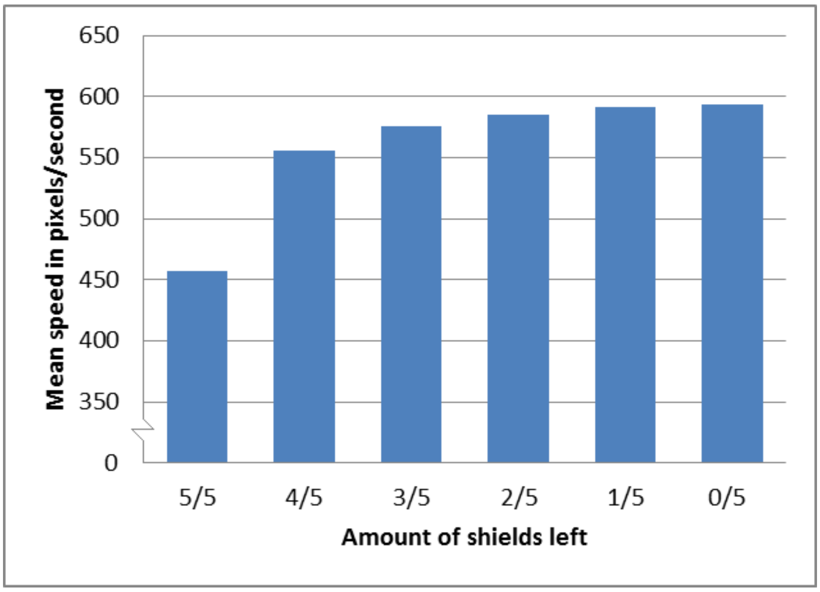

Figure 17. Mean speed within conditions: 5-shield condition.

Shields left	5/5	4/5	3/5	2/5	1/5	0/5
5/5		*	*	*	*	*
4/5			-	*	-	-
3/5				-	-	-
2/5					-	-
1/5						-
0/5						

Table 17. Significant differences in speed within conditions: 5-shield condition.

TTC:

The assumption of sphericity was violated, $X^2(9) = 33.662$, $p < .001$. Thus, degrees of freedom for the univariate results were corrected with the Huynh-Feldt estimates of sphericity ($\varepsilon = .829$). The multivariate tests were significant ($p = .006$). The univariate result showed a significant difference in mean TTC between the shields left within conditions, $F(3.317, 205.666) = 4.230$, $p = .005$. To view this difference in more detail, Bonferroni-corrected post-hoc pairwise comparisons were conducted. These revealed that there was significant difference in the mean speed between 4 and 2 shields left ($p = .011$), 4 and 1 shields left ($p = .038$) and 4 and 0 shields left ($p = .020$). This is illustrated in table 19.

	N	M	SD
4/5 shields	63	0,79	0,28
3/5 shields	63	0,74	0,28
2/5 shields	63	0,73	0,27
1/5 shields	63	0,71	0,24
0/5 shields	63	0,70	0,28

Table 18. Descriptive statistics within 5-shields condition: TTC.

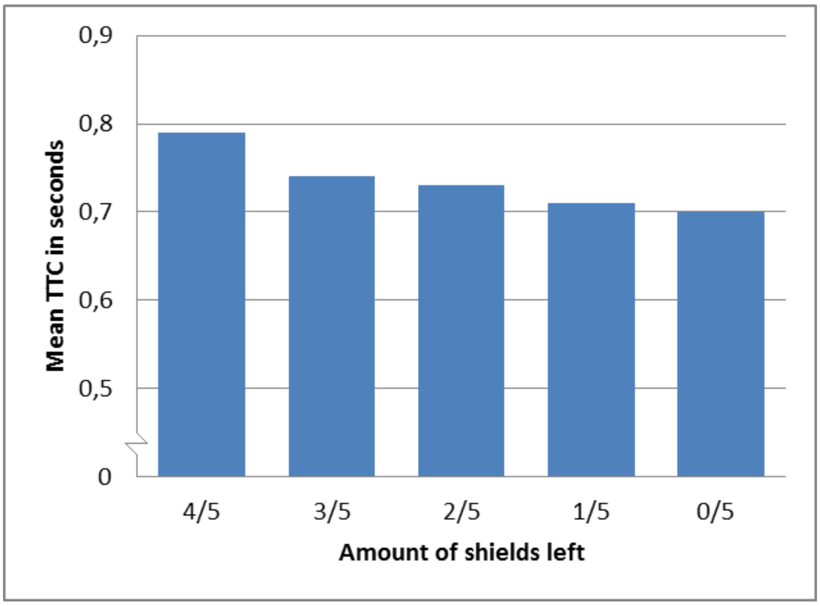

Figure 18. Mean TTC within conditions: 5-shield condition.

Shields left	4/5	3/5	2/5	1/5	0/5
4/5		-	*	*	*
3/5			-	-	-
2/5				-	-
1/5					-
0/5					

Table 19. Significant differences in TTC within conditions: 5-shield condition.

DCM:

The assumption of sphericity was met, $X^2(9) = 16.247$, $p = .061$. The multivariate tests were not significant ($p = .812$). The univariate result showed no significant difference in mean DCM between the shields left within conditions, $F(4, 248) = 0.455$, $p = .768$. Thus, no post-hoc tests were conducted.

	N	M	SD
4/5 shields	63	207,89	13,90
3/5 shields	63	205,25	18,84
2/5 shields	63	207,48	12,93
1/5 shields	63	205,43	14,42
0/5 shields	63	205,65	17,47

Table 20. Descriptive statistics within 5-shields condition: DCM.

3.1.4 Hypothesis 1b: Effectiveness of risk-taking strategies

Spearman's correlation and Kendall's tau-b were used to assess the relation between risk-taking parameters speed, TTC and DCM and the total score. This was done to see if a particular risk strategy would yield the most points and thus be the most effective.

The relation between the risk-taking parameters and score in points is visualized for the 10th to 100th percentile in figure 19.

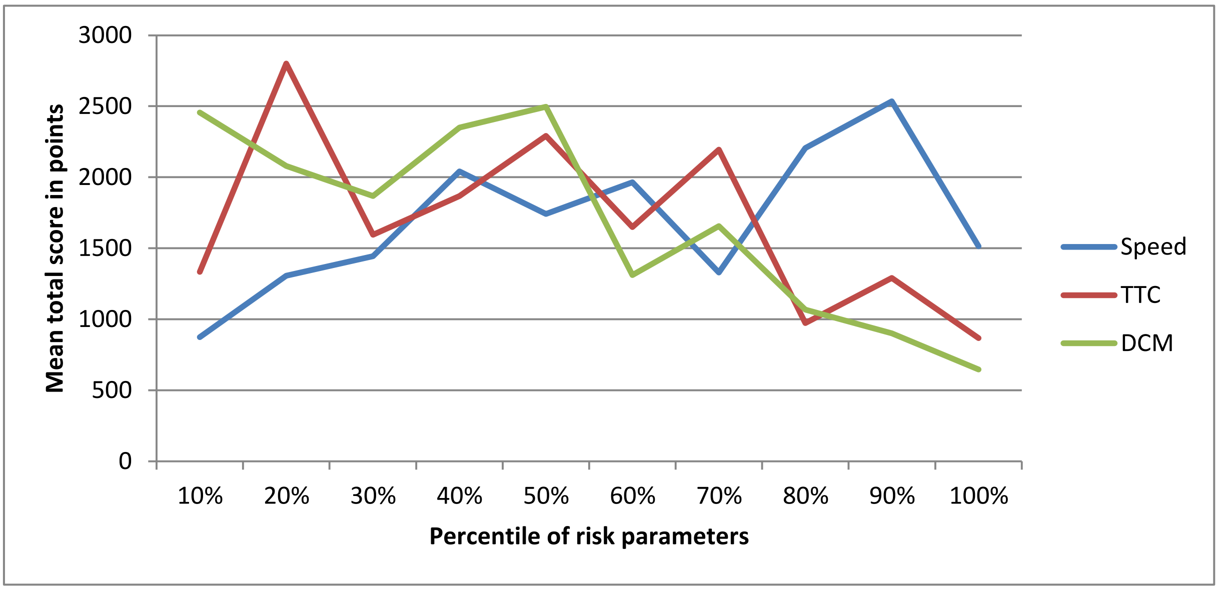

Figure 19. Mean total score for each percentile of the risk parameters. The general directions of the lines suggest a positive association between speed and total score and a negative association between TTC and DCM and total score.

Speed:

There was a weak positive association between mean speed and total score, τb = .195. This association was statistically significant (p = .018). In other words: higher speed and higher scores have a light tendency to appear together. This relation is shown in figure 20.

		Mean speed	Mean TTC
Total score	Correlation Coefficient	,195	-,191
	Sig. (2-tailed)	0,018	0,02
	N	69	69

Table 21. Kendall's tau-b results: mean speed and TTC x total score.

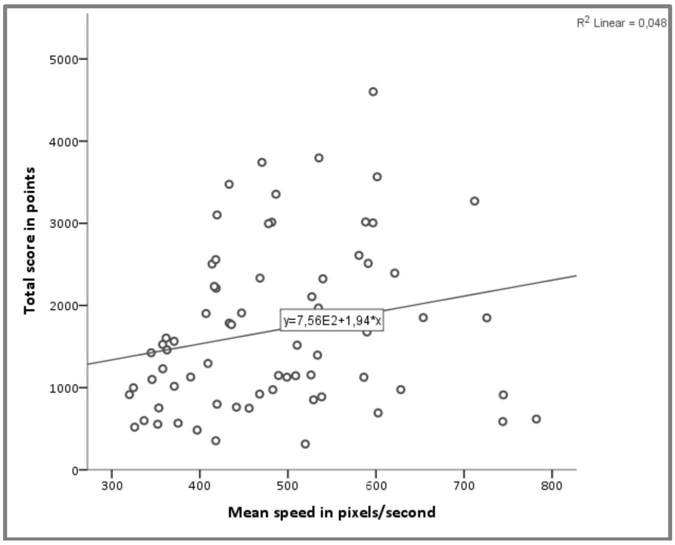

Figure 20. Scatterplot: mean speed x total score

TTC:

There was a weak, statistically significant negative association between mean TTC and total score, $\tau b = -.191$, $p = .02$. This points to a small relation between lower time to collision and higher scores. This can be seen in figure 21.

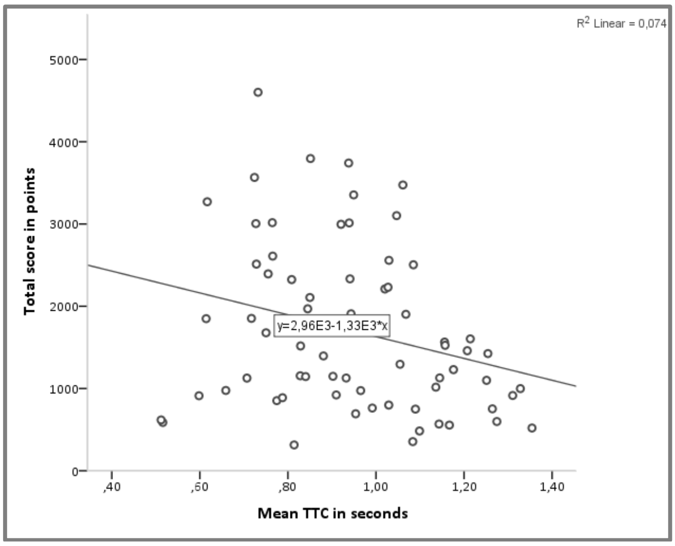

Figure 21. Scatterplot: mean TTC x total score

DCM:

According to Spearman's test, there was a moderate negative correlation between mean DCM and total score, $r_s(67) = -.620$, $p = .000$. This means a low distance to the closest meteor is associated with a high total score. This association is represented visually in figure 22.

		Mean DCM
Total score	Correlation Coefficient	-,620
	Sig. (2-tailed)	0,00
	N	69

Table 22. Spearman's correlation results: mean DCM x total score.

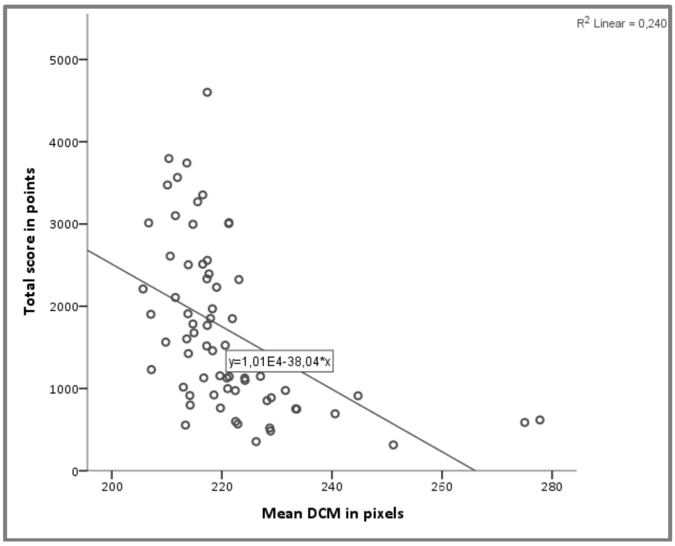

Figure 22. Scatterplot: mean DCM x total score.

An inspection of the mean points per second for each condition confirms this relation. Figure 23 shows that more points are generally gained in conditions with more shields (and thus higher expected risk-taking). The only condition not compliant with this trend is the 0 shield condition.

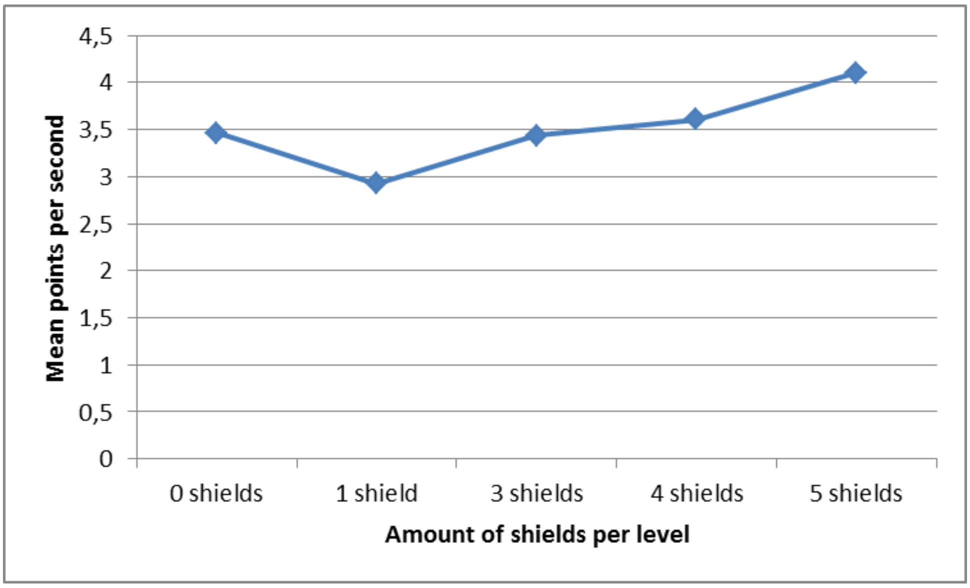

Figure 23. Mean points per second for each shield condition.

3.2 Hypothesis 2: Music exposure and risk-taking

3.2.1 Assumptions tests

Hypothesis 2a: One-way MANOVA

The relation between weekly hours of listening to music and risk-taking parameters speed, TTC and DCM was to be tested with a MANOVA, because of the multiple dependent variables. While testing the assumptions, it became clear that MANOVA could not be used to test an effect on all risk-taking parameters at once, as many were not met. Testing of the assumptions went as follows:

There were more than two dependent variables, all measured on a continuous level (speed, TTC and DCM). The independent variable, hours of music per week, consists of 5 independent categorical groups. Independence of observations is present. Sample size in all groups was sufficient: each group had more cases than the amount of dependent variables, 3 in this case. Outliers were tested for using boxplots. Some outliers were found for DCM, but these will be kept included due to them being an important part of the researched phenomenon.

Multivariate normality could not directly be tested for using SPSS Statistics so separate Shapiro Wilk tests, along with Q-Q plots were ran for an approximation. Normality of the dependent variables was tested on each level of the independent variable. For speed, scores were not normally distributed in the 'more than 6 hours'-group (p = .020). DCM was not normally distributed in groups '4 – 6 hours' (p = .001) and 'more than 6 hours' (p = .000). Results for TTC indicated normality on all levels, as well as all remaining levels of speed and DCM. Because of a MANOVA's robustness to non-normality, all data would be used in the analysis, although it would be noted along with the results.

The assumption of multicollinearity was tested using Pearson correlation between the dependent variables. The test showed that there was a moderate relation between speed and DCM (r = .454, p = .000) and TTC and DCM (r = -.363, p = .002). A high correlation was found between speed and TTC (r = -.961, p = 0.000). This means that either speed or TTC would need to be removed from the analysis for a meaningful result.

For the assumption of linearity, the data was split on the independent variable, 'hours of music', and scatterplot matrices were run for each of the dependent variables. The only linear relation was found between speed and TTC, on all of the independent variable levels. No linear relation of any kind could be found between TTC and DCM, on all levels of the independent variable.

Due to violations of multicollinearity and linearity, all 3 dependent variables were tested using separate ANOVAs.

Hypothesis 2b: One-way ANCOVA

Separate ANCOVAs were planned to test the relation between weekly hours of listening to music and risk-taking parameters speed, TTC and DCM with 'Emotional contagion score' (EC) as covariate. Assumption tests showed that Emotional Contagion was not suitable to use as a covariate of risk-taking.

The dependent variables and the covariate were continuous, and the independent variable was categorical with more than one independent group. Observations were independent.

For testing the linear relation between the covariate and the dependent variables, grouped scatterplots were used. The scatterplots showed no linearity between the covariate and the dependent variable on any level of the independent variable. The assumption of homogeneity of regression slopes was met, $F(3, 61) = 1.955$, $p = .130$.

As for normality of within-group residuals, data for the dependent variables speed, TTC and DCM is normally distributed on almost all levels of the independent variables, according to the Shapiro Wilk test. Groups not normally distributed were 'more than 6 hours' ($p = .039$) for speed and '4 – 6 hours' ($p = .001$) and 'more than 6 hours' ($p = .000$) for DCM. Because of ANCOVA's robustness to non-normality, all data would be included in the analysis, although non-normality would be noted. Scatterplots were used to check on homoscedasticity. The scatterplots showed homoscedasticity in each level of the independent variable. Outliers were not removed as these are relevant to the subject being researched.

Due to violations in linearity between the covariate and the dependent variables, ANCOVA could not be used. Instead, the data was split into two groups by Emotional Contagion score resulting in the groups 'low susceptibility' and 'high susceptibility'. The relation between hours of music per week and the risk parameters will be tested separately for each group using ANOVAs.

Assumptions overlapping for ANCOVA and ANOVA were met and will be discussed in the next section.

Hypothesis 2b: Independent samples t-test

Independent samples t-tests were conducted to further assess the difference in risk-taking between low susceptibility and high susceptibility, which was especially large in the ‘4 – 6 hours’ and ‘more than 6 hours’ groups.

The first three assumptions for this test were met, as the dependent variables were continuous, the independent variable is categorical with two groups and there was independence of observations.

Outliers were checked with boxplots and found, but because these were minimal and relevant to the question at hand, they were mostly not removed. For the ‘more than 6 hours’ group for DCM, there were too many for the data to be usable. This group was not analyzed. Outliers were kept in mind while reporting the results.

The dependent variables were almost all normally distributed, as assessed by Shapiro-Wilk’s test. Only the ‘Low susceptibility’ group for DCM violated this assumption, $p = .013$. The group was still included in the analysis, as the independent samples t-test is robust to non-normality,

The assumption of homogeneity of variance, as according to Levene’s test, was addressed and reported with the test results.

Hypotheses 2a and 2b: One-way ANOVA

The dependent variables (speed, TTC and DCM) were continuous. The independent variable, 'hours of music', was categorical with 4 groups. Observations were independent.

Outliers were assessed using boxplots, but not removed as they are relevant to the researched subject. This will be kept in mind when reporting the results.

The assumption of normality was assessed using Shapiro Wilk's test. The dependent variables were approximately normally distributed for most groups. Groups with non-normality were 'more than 6 hours' ($p = .020$) for speed and '4 – 6 hours' ($p = .001$) and 'more than 6 hours' ($p = .000$) for DCM.

For homogeneity of variances, Levene's test was run and reported during the analyses.

3.2.2 Hypothesis 2a: The relation between weekly listening hours and risk-taking

Three one-way ANOVAs were used to compare weekly hours of listening to music (a categorical variable with 5 groups) and preference in musical genre (a categorical variable with 4 groups) to risk-taking parameters speed, TTC and DCM.

Speed:

The result of Levene's test showed that there was homogeneity of variances (p = .869). The results showed no significant difference in mean speed between the groups, $F(3, 65)$ = .343, p = 794. Despite not being significant, mean speed shows a peak for the ' 4-6 hours' – group. No post-hoc tests were performed.

	N	M	SD
0.5-2 hours	8	484,83	100,45
2-4 hours	12	476,92	115,24
4-6 hours	14	513,06	102,49
> 6 hours	35	477,73	122,09

Table 23. Descriptive statistics weekly hours of listening to music: speed.

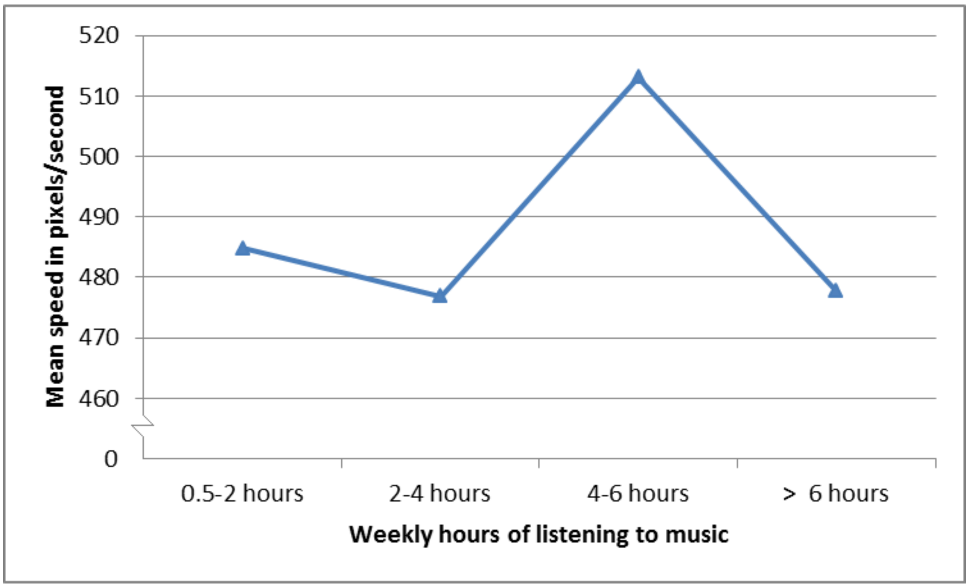

Figure 24. Illustration of mean speed for different amounts of weekly music listening.

TTC:

Levene's test showed that there was homogeneity of variances (p = .946). Although a similar peak in risky behavior as to speed can be observed in figure 25, the result showed no statistically significant difference in TTC between the groups, $F(3, 65) = .590$, $p = .624$. No post-hoc tests were performed.

	N	M	SD
0.5-2 hours	8	0,93	0,19
2-4 hours	12	0,97	0,22
4-6 hours	14	0,89	0,19
> 6 hours	35	0,97	0,21

Table 24. Descriptive statistics weekly hours of listening to music: TTC.

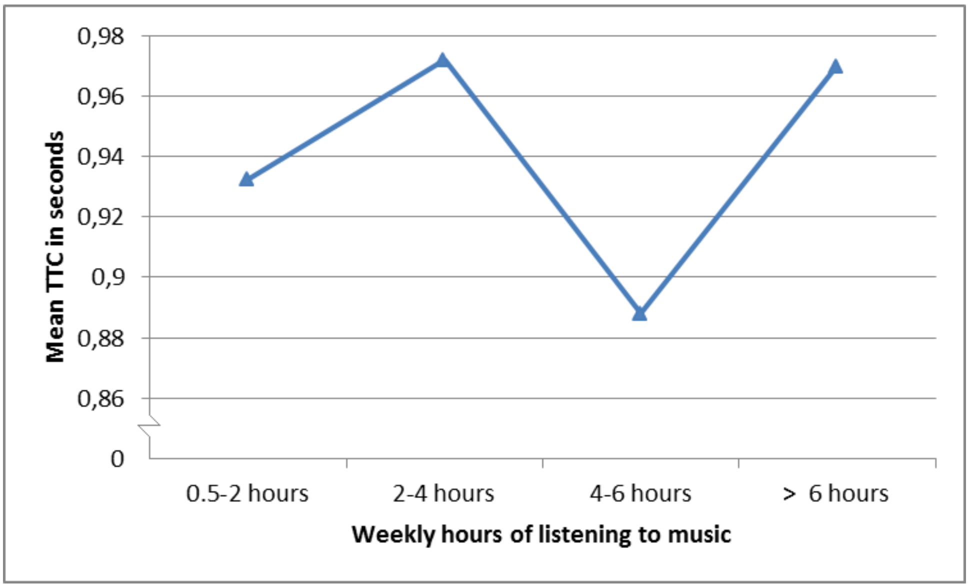

Figure 25. Illustration of mean TTC for different amounts of weekly music listening.

DCM:

The assumption of homogeneity of variances was met (p = .730). The results showed no significant difference in mean DCM between the groups, $F(3, 65) = .670$, p = 574. No post-hoc tests were performed.

	N	M	SD
0.5-2 hours	8	218,05	6,84
2-4 hours	12	220,03	9,42
4-6 hours	14	225,39	16,08
> 6 hours	35	221,08	13,62

Table 25. Descriptive statistics weekly hours of listening to music: DCM.

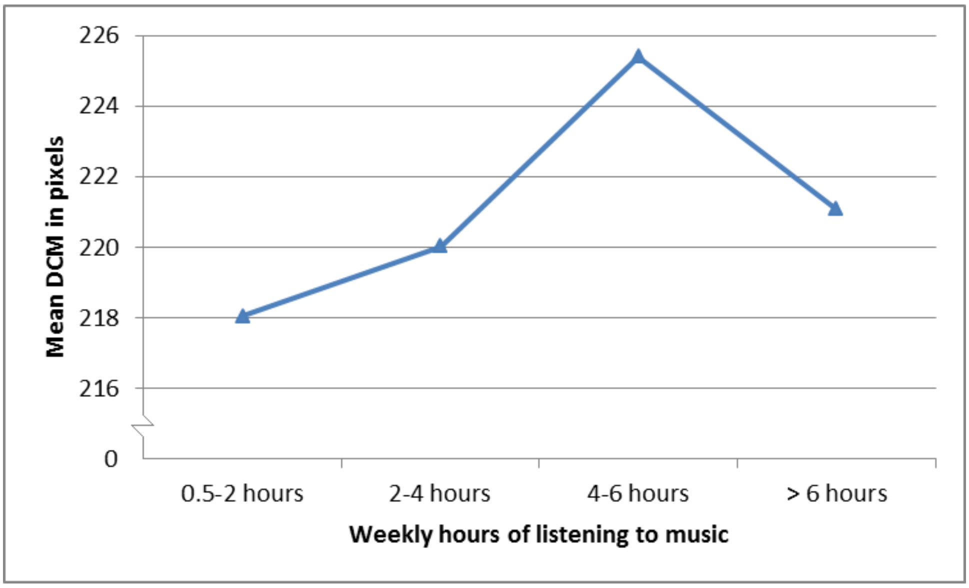

Figure 26. Illustration of mean DCM for different amounts of weekly music listening.

3.2.3 Hypothesis 2b: Emotional susceptibility in listening hours and risk-taking

To assess the effect of listening hours on risk behavior for different emotional susceptibility groups, 'split file' was used. For the groups, emotional contagion total scores were to be split between 'low' and 'high' for total susceptibility. ANOVAs were planned for each group separately. Splitting the total EC scores down the middle resulted in two unbalanced parts, in which not all 'listening hours' groups were represented (table 26). With this division, the ANOVAs could not be representatively performed.

EC - Total		Emotional susceptibility group	
		Low	High
Listening hours	0.5-2 hours	2	6
	2-4 hours	0	12
	4-6 hours	0	14
	> 6 hours	3	32
Total		5	64

Table 26. Number of participants per group for total susceptibility,split into 'Low'-'High'.

To be able to analyze further, one of the EC subscores, Fear, was used and split through the middle into 'low' and 'high' groups, resulting in more equal, analyzable groups (figure 27 and table 27).

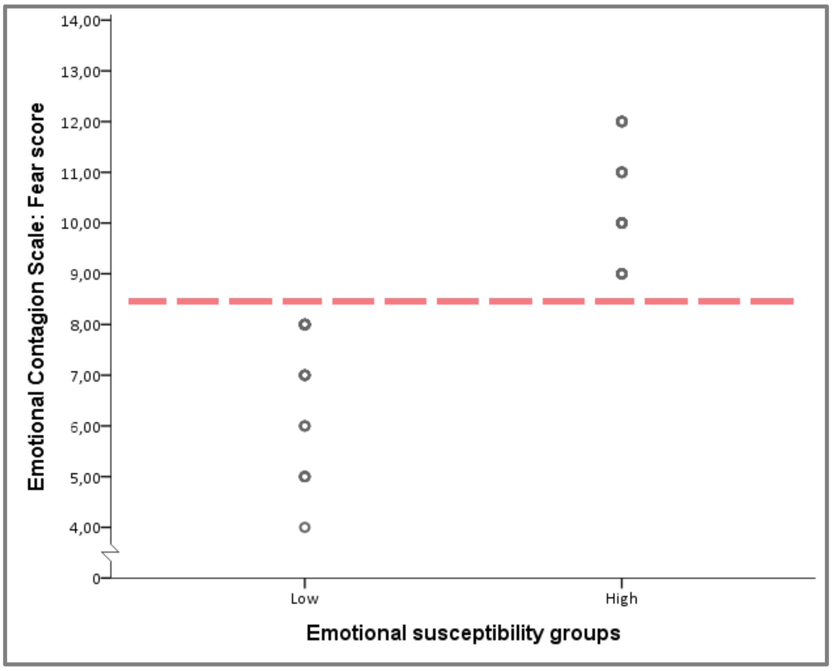

Figure 27. Division into ‘low’ and ‘high’ susceptibility groups

EC - Fear		Emotional susceptibility group	
		Low	High
Listening hours	0.5-2 hours	4	4
	2-4 hours	5	7
	4-6 hours	9	5
	> 6 hours	14	21
Total		32	37

Table 27. Number of participants per group for fear susceptibility, split into 'Low'-'High'.

Speed:

Low susceptibility to fear group:

Levene's test showed that the assumption of homogeneity of variances was met (p = .053). The ANOVA results showed that there was no statistically significant difference in mean speed between the groups, $F(3, 28) = 2.745$, p = .062. Therefore, no post-hoc tests were consulted.

High susceptibility to fear group:

The assumption of homogeneity was met, according to Levene's test (p = .337). The results showed no statistically significant difference between the groups, $F(3, 33) = .570$, p = .638. No further analyses were pursued.

	Low fear susceptibility			High fear susceptibility		
	N	M	SD	N	M	SD
0.5-2 hours	4	484,47	124,71	4	485,19	89,39
2-4 hours	5	490,27	156,99	7	467,38	87,54
4-6 hours	9	548,18	99,25	5	449,83	81,69
> 6 hours	14	422,74	75,86	21	514,39	134,44
Total	32	476,29	111,69	37	493,62	115,69

Table 28. Descriptive statistics of speed over different 'hours of listening'-groups, divided by low and high susceptibility to fear.

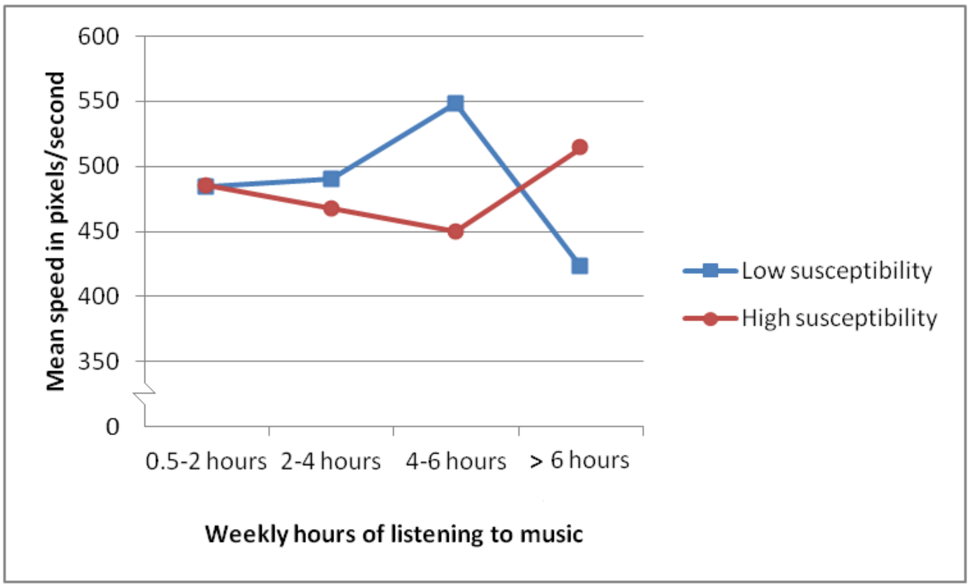

Figure 28. Illustration of mean speed for different amounts of weekly music. The different lines represent the two fear susceptibility groups.

TTC:

Low susceptibility to fear group:

Levene's test showed homogeneity of variances (p = .081). The ANOVA results showed a significant difference in mean TTC between the groups, $F(3, 28) = 3.204$, p = .038. A post-hoc Tukey test was performed for a more detailed view of the between-group differences. There was a significant difference between the '4 to 6 hours' and 'more than 6 hours' groups (p = .024).

High susceptibility to fear group:

The assumption of homogeneity of variances was met (p = .739). The results showed no significant difference in mean TTC between the groups, $F(3, 33) = .477$, p = .700. No post-hoc tests were performed.

	Low susceptibility			High susceptibility		
	N	M	SD	N	M	SD
0.5-2 hours	4	0,93	0,24	4	0,94	0,17
2-4 hours	5	0,95	0,27	7	0,99	0,19
4-6 hours	9	0,83	0,17	5	1,00	0,19
> 6 hours	14	1,07	0,15	21	0,90	0,22
Total	32	0,97	0,21	37	0,93	0,20

Table 29. Descriptive statistics for TTC over different 'hours of listening'-groups, divided by low and high susceptibility to fear.

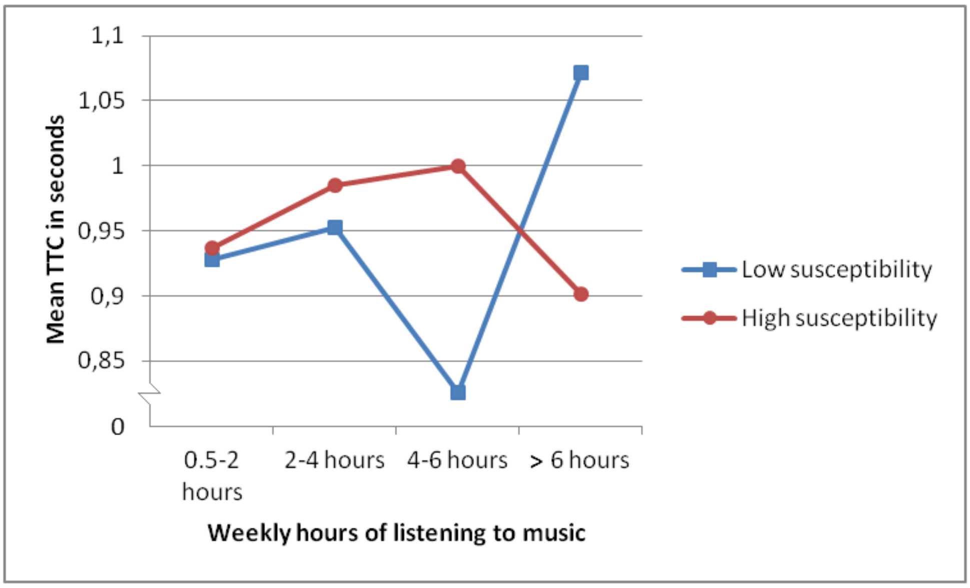

Figure 29. Illustration of mean TTC for different amounts of weekly music. The different lines represent the two fear susceptibility groups.

DCM:

Low susceptibility to fear group:

The homogeneity of variance assumption was met (p = .330). The test results showed no significant difference in mean DCM between the 'hours of music' groups, $F(3, 28) = 1.204$, p = .326. No further testing was conducted.

High susceptibility to fear group:

Levene's test showed that the assumption of homogeneity was met (p = .357). No significant difference was found between the groups, $F(3, 33) = .294$, p = .829. With the lack of significant group differences, no post-hoc test were performed.

	Low susceptibility			High susceptibility		
	N	M	SD	N	M	SD
0.5-2 hours	4	219,38	9,40	4	216,72	3,99
2-4 hours	5	217,78	10,37	7	221,63	9,16
4-6 hours	9	227,56	19,84	5	221,47	4,70
> 6 hours	14	217,58	8,21	21	223,40	16,05
Total	32	220,65	13,11	37	222,08	12,85

Table 30. Descriptive statistics for DCM over different 'hours of listening'-groups, divided by low and high susceptibility to fear.

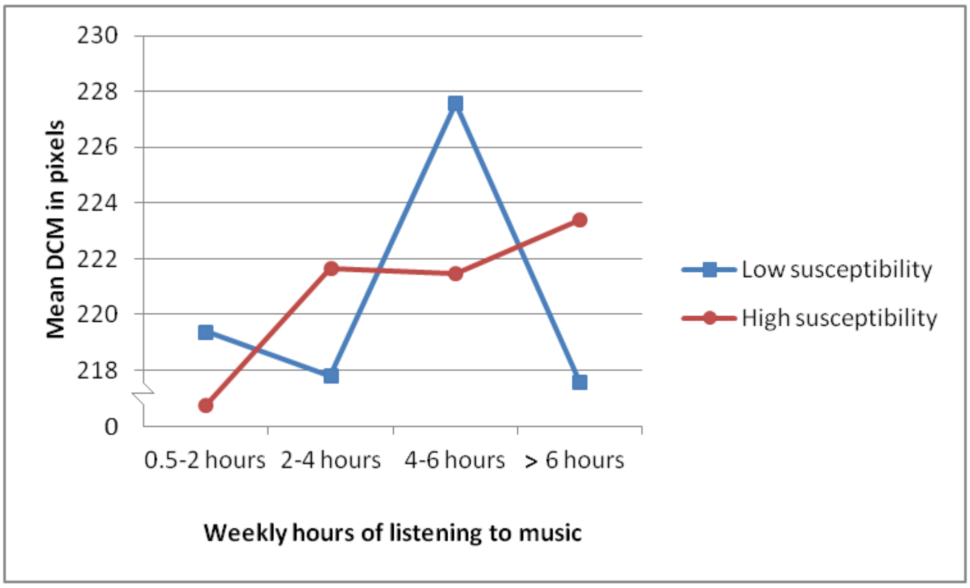

Figure 30. Means plot for DCM for different amounts of weekly music. The different lines represent the two fear susceptibility groups.

Due to the large differences between the low and high susceptibility to fear groups for '4-6 hours' and 'More than 6 hours', paired samples t-tests were conducted to assess whether these differences were significant.

'4 – 6 hours per week' group

Speed:

The assumption of homogeneity of variances was met, according to Levene's test (p = .940). Mean speed in the 'low susceptibility' group was 98,35 pixels per second higher than in the 'high susceptibility' group. However, this difference was not statistically significant, t(12) = 1.881, p = .085.

TTC:

Levene's test showed that there was homogeneity of variances for TTC as well (p = .672). The mean time to collision in the 'low susceptibility' group was 0.17 seconds lower than in the 'high susceptibility' group. Although many times larger than differences in other listening hours-groups, this difference was not significant, t(12) = -1.806, p = .096.

DCM:

There was homogeneity of variances (p = .199). The mean distance to the closest meteor in the 'low susceptibility' group was 6.09 pixels higher than in the 'high susceptibility' group. This difference was not significant, t(12) = .664, p = .519.

'More than 6 hours per week' group

Speed:

The assumption of homogeneity of variances was met (p = .053). Mean speed in the 'low susceptibility' group was 91.66 seconds lower than in the 'high susceptibility' group, a significant difference, t(33) = -2.31, p = .027. Although significant, this outcome is opposite to the general trend where speed tends to be higher in the 'low susceptibility' group. This will be kept in mind.

TTC:

There was homogeneity of variances, according to Levene's test (p = .174). Mean TTC in the 'low susceptibility' group was 0.17 seconds higher than in the 'high susceptibility' group. Again, this difference is reversed from the regular trend but significant, t(33) = 2.484, p = .018.

3.3 Hypothesis 3: Music preference and risk-taking

3.3.1 Assumptions tests

The relation between music preference and risk-taking parameters speed, TTC and DCM was assessed using separate ANOVAs. For hypothesis 3a, the ANOVA was used directly. For hypothesis 3b, the file was split on 'low' and 'high' emotional susceptibility. The effect will be analyzed separately for these groups, allowing for comparison.

Hypotheses 3a and 3b: ANOVA

The dependent variables (speed, TTC and DCM) were continuous. The independent variable, 'music preference', was categorical with 4 groups ('pop', 'rock', 'electronic' and 'other'). Observations were independent.

Outliers were assessed using boxplots, but not removed as they are relevant to the researched subject. This will be kept in mind when reporting the results.

The assumption of normality was assessed using Shapiro Wilk's test. The dependent variables were approximately normally distributed for most groups. Groups with non-normality were 'pop' (p = .000) and 'electronic' (p = .000) for DCM. Because of an ANOVA's robustness to non-normality, the groups were used in the analysis. The violations were kept in mind when reporting and concluding.

For homogeneity of variances, Levene's test was run and reported during the analyses.

Hypothesis 3b: Independent samples t-test

To look into the large difference between the two susceptibility groups in the ‘Electronic’ preference, three independent samples t-tests were used.

The assumptions of continuous dependent variables, categorical independent variables and independence of observations were met.

Boxplots were used to check for outliers. Outliers were only found for the DCM groups. Due to their relative amount and size, the data is not completely representative to use. The data was also checked for an approximately normal distribution. All groups were normally distributed according to Shapiro-Wilk’s test, except for those in DCM, p = .006 (‘low’), p = .001 (‘high’). Although the test is robust to non-normality, the combination with outliers made the DCM groups unfit to use for analysis.

Levene’s test for homogeneity of variances was run during the analyses.

3.3.2 Hypothesis 3a: The relation between music preference and risk-taking

Speed:

The assumption of homogeneity of variances was met, as shown by Levene's test (p = .909). The results show no statistically significant difference in speed between the music preference groups, $F(3, 65) = .665$, p = .576. No post-hoc tests were performed.

	N	M	SD
Pop	27	484,85	114,95
Rock	9	440,72	91,47
Electronic	10	485,30	136,01
Other	23	504,12	110,94

Table 31. Descriptive statistics of speed for different music preferene groups.

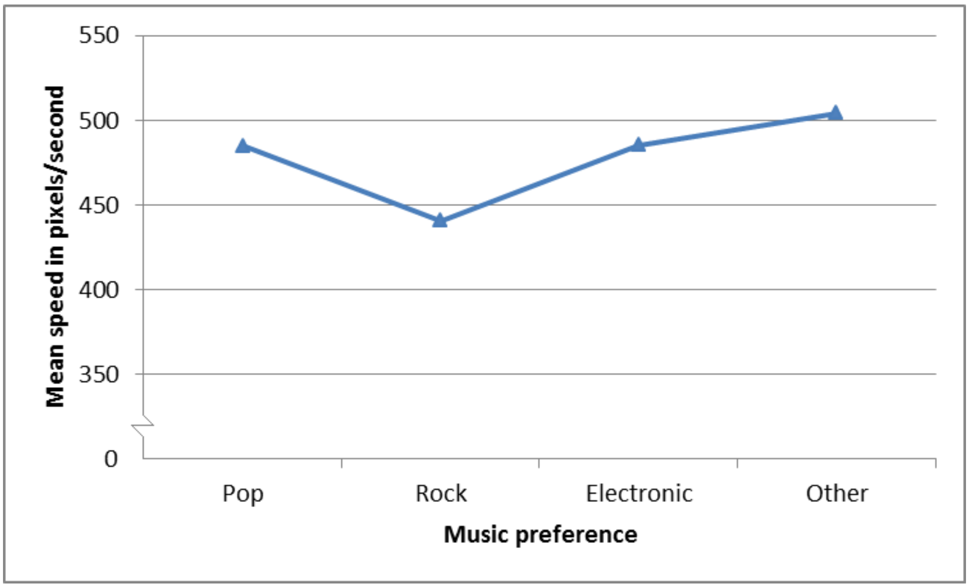

Figure 31. Illustration of mean speed for musical preference groups Pop, Rock, Electronic and Other.

TTC:

The assumption of homogeneity of variances was met, as shown by Levene's test (p = .945). The results show no statistically significant difference in TTC between the groups, $F(3, 65)$ = .520, p = .670. No further testing was pursued.

	N	M	SD
Pop	27	0,95	0,21
Rock	9	1,02	0,19
Electronic	10	0,96	0,24
Other	23	0,92	0,20

Table 32. Descriptive statistics for TTC over different music preference groups.

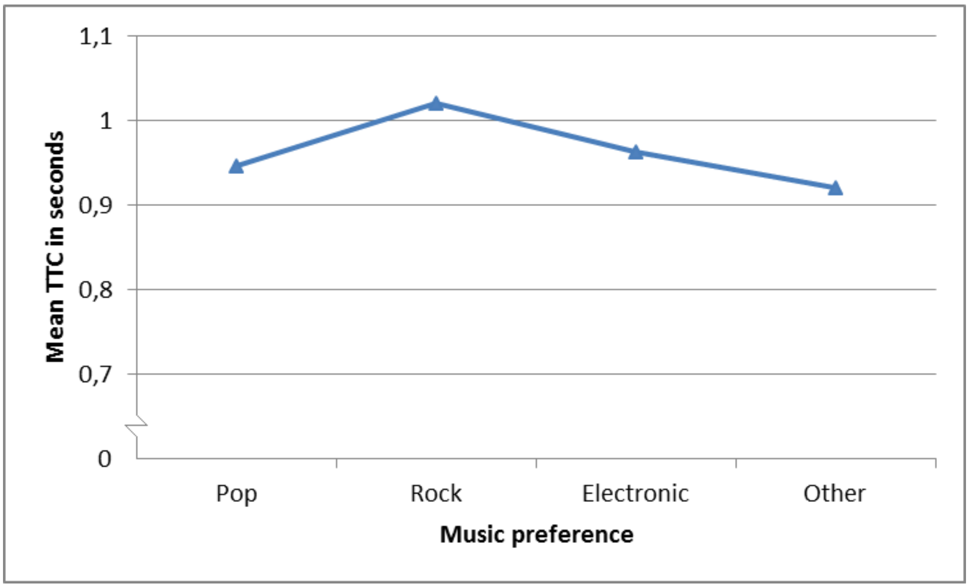

Figure 32. Means plot of TTC for different musical preference groups.

DCM:

The assumption of homogeneity of variances was met, as shown by Levene's test (p = .239). There was no statistically significant difference in DCM between the groups, $F(3, 65)$ = .934, p = .430. No post-hoc tests were performed.

	N	M	SD
Pop	27	221,67	12,17
Rock	9	214,90	5,45
Electronic	10	223,17	19,69
Other	23	222,93	12,19

Table 33. Descriptive statistics for DCM for different musical preferences.

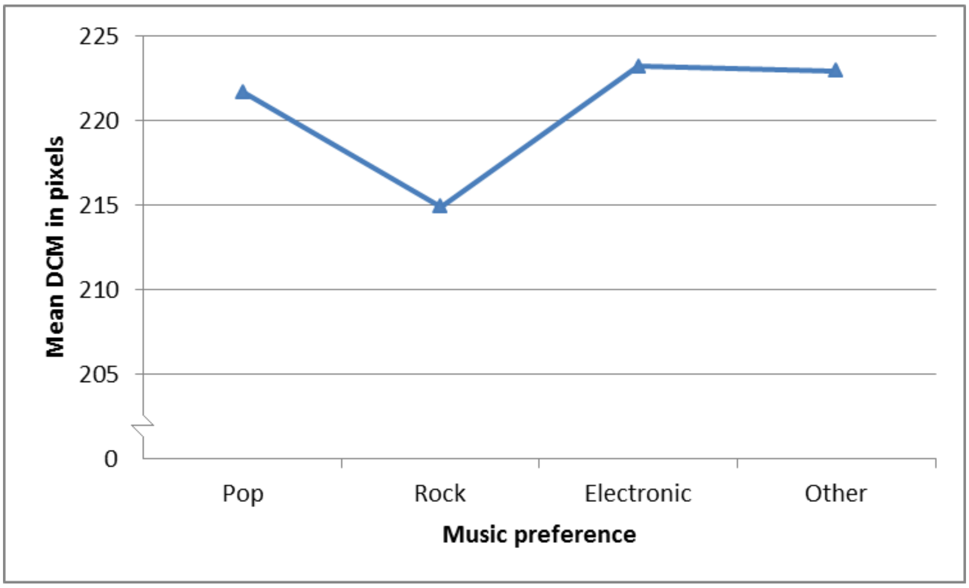

Figure 33. Illustration of mean DCM for different musical preference groups.

Despite a lack of evidence for a relation to risk, players' in-game success for each group was visually represented for further insight (Figure 34). When looking at performance, preference groups Rock and Electronic had a higher than average score in points per person. For Pop, this was below average. This suggests that the relation between risk and performance is not a direct one.

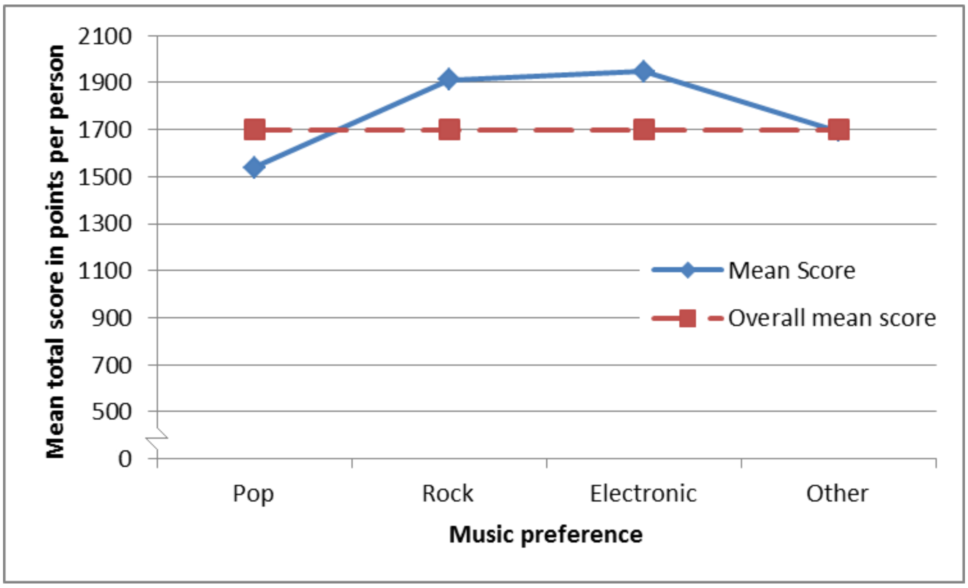

Figure 34. Mean score in points for each preference group.

3.3.3 Hypothesis 3b: Emotional susceptibility in music preference and risk-taking

To test the relation between music preference and risk-taking for different emotional susceptibility groups, three ANOVAs were conducted. This was done with a data file split between 'low fear susceptibility' and 'high fear susceptibility'.

Speed:

Low group:

The assumption of homogeneity of variances was met, as shown by Levene's test (p = .157). The results showed no significant difference in mean speed between the music preference groups, $F(3, 28)$ = .1204, p = .326. Thus, no post-hoc tests were performed.

High group:

Levene's test showed that the assumption of homogeneity of variances was met (p = .642). No significant difference between the groups could be found, $F(3, 33)$ = .985, p = .412. Group differences were not analyzed further.

	Low susceptibility			High susceptibility		
	N	M	SD	N	M	SD
Pop	11	485,87	130,83	16	484,13	107,20
Rock	5	445,03	77,47	4	435,32	119,33
Electronic	5	407,03	64,44	5	563,58	148,84
Other	11	512,38	114,37	12	496,55	112,23
Total	32	476,29	111,70	37	493,62	115,70

Table 34. Descriptive statistics for speed over different music preference groups, divided by low and high susceptibility to fear.

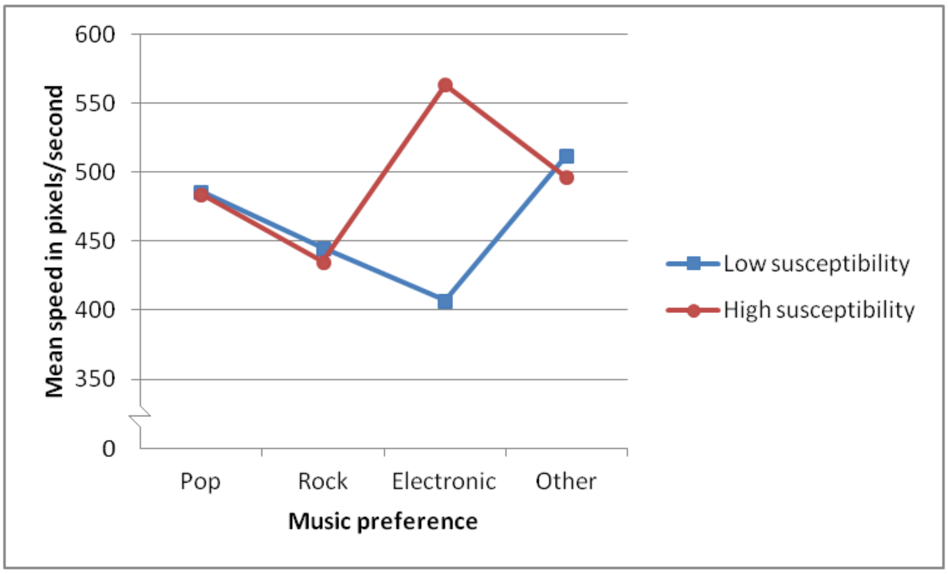

Figure 35. Means plot for speed over different music preferences. The lines represent the two fear susceptibility groups.

TTC:

Low group:

The assumption of homogeneity of variances was met (p = .551). The results pointed to no significant differences in mean TTC between the groups, $F(3, 28) = 1.158$, p = .343. No further testing was conducted.

High group:

Levene's test showed that there was homogeneity of variances (p = .982). The results showed no significant differences between the groups, $F(3, 33) = 1.011$, p = .400. No post-hoc tests were performed.

	Low susceptibility			High susceptibility		
	N	M	SD	N	M	SD
Pop	11	0,94	0,23	16	0,95	0,20
Rock	5	0,99	0,17	4	1,05	0,24
Electronic	5	1,11	0,18	5	0,82	0,20
Other	11	0,91	0,21	12	0,93	0,20
Total	32	0,97	0,21	37	0,93	0,20

Table 35. Descriptive statistics for TTC over different music preference groups, divided by low and high susceptibility to fear.

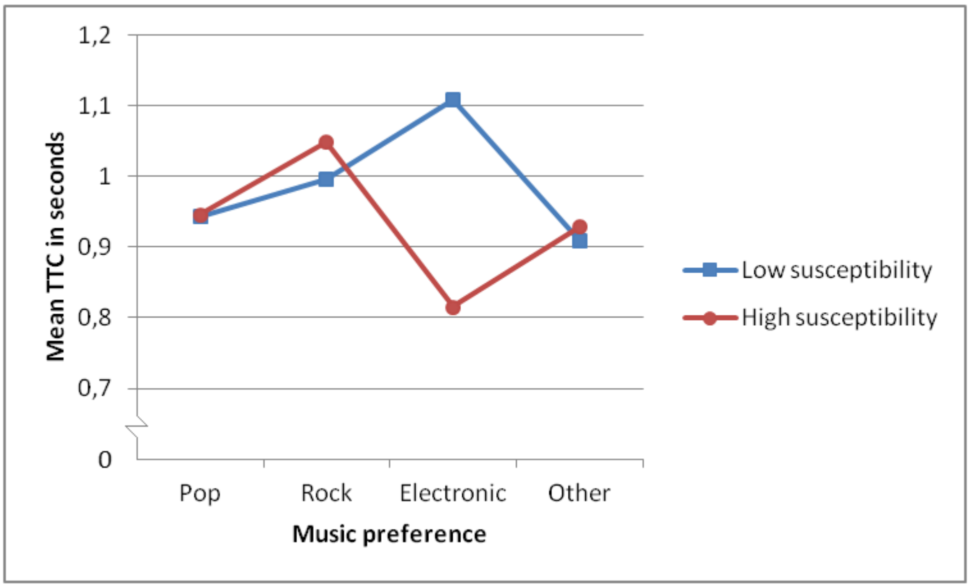

Figure 36. Illustration of mean TTC for different music preferences. The lines represent the two fear susceptibility groups.

DCM:

Low group:

The assumption of homogeneity of variances was met (p = .285). There were no statistically significant differences in mean DCM between the music preference groups, $F(3, 28) = .926$, p = .441. Therefore, no further testing was pursued.

High group:

Levene's test showed that the assumption of homogeneity of variances was violated (p = .003). Therefore, the Welch ANOVA results were used. These showed no significant differences in mean DCM between the groups, Welch's $F(3, 8.953) = .761$, p = .544. No follow-up tests were performed.

	Low susceptibility			High susceptibility		
	N	M	SD	N	M	SD
Pop	11	224,06	18,27	16	220,02	5,29
Rock	5	213,25	5,37	4	216,96	5,52
Electronic	5	217,32	6,47	5	229,02	27,29
Other	11	222,13	11,10	12	223,66	13,56
Total	32	220,65	13,12	37	222,09	12,85

Table 36. Descriptive statistics for DCM over different music preference groups, divided by low and high susceptibility to fear.

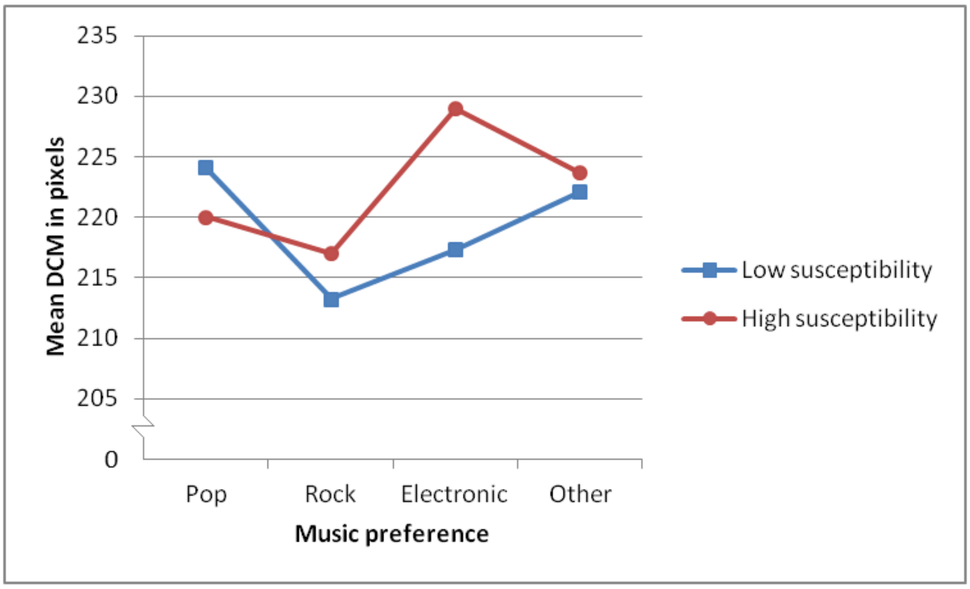

Figure 37. Means plot for DCM over different music preferences. The lines represent the two fear susceptibility groups.

To further inspect the large difference between low and high susceptibility in the 'Electronic' group specifically, independent samples t-tests were conducted with a split file. The DCM groups were not included in the analysis due to assumption issues.

Speed:

Levene's test showed that the assumption of homogeneity of variances was violated (p = .034). Therefore, corrected Welch t-values were used. Mean speed for the 'low susceptibility' group was 156.55 seconds lower than in the 'high susceptibility' group. This difference was not significant, t(5.449) = -2.158, p = .079.

TTC:

For TTC, the assumption of homogeneity was met (p = .703). Mean time to collision in the 'low susceptibility' group was 0.29 seconds higher than in the 'high susceptibility' group. In terms of TTC, the 'low susceptibility' group flew significantly safer, t(8) = 2.47, p = .039.

4. Discussion

4.1 Hypothesis 1

4.1.1 Between conditions

To answer the question of whether risk homeostasis could be found in The Spaceship Game, players' risk-taking behavior was compared between conditions (0, 1, 3, 4 or 5 shields) on multiple levels. For this, Repeated Measures ANOVAs were used, with each condition as a time point.

Significant differences were found between the different conditions for each of the risk parameters speed, time to collision (TTC) and distance to the closest meteor (DCM). However, this difference was not significant between all adjacent conditions, for example between 3 and 4 shields. Significant differences were generally found between the amounts of shields that were more than 1 place apart. A possible explanation for this could be that the different levels of risk were unconsciously perceived as too close together. When starting with multiple shields, the difference between 3 or 4 shields might not be a noticeable difference in risk. This effect could be researched further with either more pronounced differences in shields (and thus protection), or an additional representation of protection that is communicated to the player more directly (such as the visual state and type of the spaceship, or the human passenger being directly exposed at the last shield).

When assessing the group means visually, a steady linear relation can be found between the conditions. When a session was started with more shields, the amount of risk taken went up as well: speed was higher, TTC was lower and DCM was lower as well.

These results were mostly in accordance with the hypothesis. Although group differences were not always significant, there was a steady line of development showing that higher protection levels did, in fact, lead to more risk-taking.

4.1.2 Within conditions

Significant risk homeostasis effects were found between the different conditions: higher levels of protection led to more risk-taking behavior. The question was: can this effect be observed on a more specific level as well? To this end, multiple Repeated Measures ANOVAs were conducted again, this time for each shield left during a condition.

Although significant differences (mainly in speed and TTC) were found between different amounts of shields left within the conditions, the direction of this effect was often in the opposite way: with more of their shields destroyed, participants generally flew the rocket more dangerously. A possible factor in this could be player habituation: on one side, 'starting up' and finding their way around the controls and, on the other side, less involvement due to boredom. On a global scale, these maneuvers were still performed within a generally higher or lower risk-taking spectrum, as found in the 'between conditions'-results.

4.1.3 Risk Homeostasis

In this research, evidence for a risk homeostasis effect was found on a larger scale between conditions, but not within those conditions. This could be a result of lagged feedback, but the possibility also exists that on a smaller, per-second scale, several processes are at play. On a between-condition scale, the relation between more distinctive amounts of perceived risk and a stable, pronounced homeostasis effect could be incentive for further investigation.

4.1.4 Hypothesis 1b: Effectiveness of risk-taking strategies

The question if a certain amount of risk would lead to an optimal score was approached with a correlational analysis.

Weak relations to total score were found for speed and TTC. For DCM, this relation was medium-sized. All relations pointed in the same direction: A higher risk is accompanied by a higher score, especially with DCM: flying closer to the on-screen meteors often occurred together with a higher score. This trend was also found when looking at each session separately: generally, more points per second were collected in sessions with more shields. As discussed earlier, more shields per level are connected to higher risk-taking. However, as this analysis was correlational, conclusions should be drawn with caution. Further research is necessary to confidently determine a causal relationship. Another reason for a cautious approach is the interpretation of ‘success’ as a result of risk-taking. In The Spaceship Game, success is expressed as the total amount of points collected through the five sessions, in which an acceleration in flying speed meant a faster point influx. To be able to attribute real-life successful situations to the amount of risk taken, these findings need to be supplemented with various other types of concrete measures of success, suited to their respective situations.

Returning to the question at hand: the results from this research provide moderate evidence for the benefits of risk-taking, in this specific game environment.

4.2 Hypothesis 2

4.2.1 Hypothesis 2a: Music exposure and risk-taking

In hypothesis 2a, it was expected that more weekly hours of listening to music would lead to a higher amount of risk-taking. This was tested using one-way ANOVAs. No significant effects were found, but an unusual peak in risk-taking occurred in the '4 – 6 hours' group for speed and TTC. For DCM, although not significant, there was a general trend of more distance being kept with more listening hours per week. For this parameter, a peak occurred in the '4 – 6 hours' group as well, but this time for careful behavior instead of risk-taking, showing an exceptionally large distance to the closest meteor. Due to the lack of significance and small and unequal group sizes, little to no conclusions can be drawn about this group yet. A possible explanation for the lack of influence of this this type of music involvement is its passive nature. An interesting follow-up for this research would be one using music practice in addittion listening as a variable for involvement. As many effects of music on the brain are related to its practice (Corrigall et al., 2013), this active involvement with music might be more directly related to human behavior than listening only. Seeing the substantial size of the 'more than 6 hours' group, including weekly involvement hours beyond 6 separately in the future could expand the spectrum and lead to more specific results. For now, no conclusive relation between exposure to music and risk-taking can be described.

4.2.2 Hypothesis 2b: Emotional susceptibility in exposure and risk-taking

To insert the factor of emotion into the equation, hypothesis 2b describes an expected difference in the relation between listening to music and risk taking for different emotional susceptibility groups. This was tested using the same ANOVAs as with hypothesis 2a, but this time with the sample split into the 'low susceptibility to fear' and 'high susceptibility to fear' groups. Where differences between the fear susceptibility groups looked especially large, these were analyzed using independent samples t-tests.

The results were mixed: there were hardly any significant differences between 'hours of listening' for both low and high fear susceptibility. For TTC, the difference between '4 – 6 hours of listening' and '> 6 hours of listening' was significant in the 'low susceptibility' group, but not for the 'high susceptibility' group. For speed and TTC, the visual trend suggested that the group with a high susceptibility to fear generally flew a little safer than the group with a low susceptibility, until 4 – 6 hours of weekly listening. For '> 6 hours', this reversed. Although the other way around, the difference in risk between low and high susceptibility was significantly large. For DCM, the susceptibility groups were mixed in their relative risk-taking. The ‘high susceptibility’ group seemed to show a slowly developing trend of keeping more distance when more weekly listening hours were involved, but the ‘low susceptibility’ group revealed no apparent logic.

From these results, no solid evidence of emotional interference on the relation between music exposure and risk can be concluded. Although possibly also due to large differences in group sizes, the inconsistent behavior of the ‘> 6 hours per week’ group calls for closer examination of risk behavior beyond 6 hours of weekly exposure.

4.3 Hypothesis 3

4.3.1 Hypothesis 3a: Music preference and risk-taking

In hypothesis 3a, it was expected that music preference makes a difference in the amount of risk taken throughout The Spaceship Game. ANOVAs were performed to assess any differences between the preference groups. The results showed no evidence of a direct connection between music preference and risk-taking. A look into the preference groups’ performance in points confirmed findings discussed in hypothesis 1b: a better performance is not necessarily due to more risk-taking.

4.3.2 Hypothesis 3b: Emotional susceptibility in preference and risk-taking

For hypothesis 3b, emotional susceptibility was expected to play a role in the relation between music preference and risk-taking. This interaction was analyzed using ANOVAs, with a sample split between ‘low’ and ‘high’ susceptibility groups. Notable differences between these groups were assessed with independent samples t-tests.

Again, there were no significant or remarkable differences between the music preferences, for neither the low nor the high fear group. For the difference between susceptibility groups, an interesting result occurred. While scores for both groups were about the same throughout the ‘Pop’, ‘Rock’ and ‘Other’ preferences, ‘Electronic’ showed a visually outstanding difference between the two on all risk parameters. For speed and TTC, electronic music lovers with high fear susceptibility flew more dangerously than those with low susceptibility. This difference was significant only for TTC. In terms of DCM, electronic music lovers with a high susceptibility kept more distance and thus flew safer than their low susceptibility counterparts. This difference was smaller than those for speed and TTC. All in all, the role of susceptibility to fear between music preference and risk-taking was not entirely supported. According to these results, music lovers fly undisturbed by their (lack of) sensitivity…except for electronic.

4.4 Conclusion

In this experiment, evidence for Wilde’s (2012) Risk Homeostasis Theory was found. The effect manifested mainly on a large, between-condition scale. No homeostasis effect was found within conditions, after losing protection. As Wilde (2012) described, lagged feedback may cause a person to be unable to reassess their risk level in an immediate situation, in this case losing a shield as a result of collision with a meteor.

Little evidence was found for a relationship between music exposure and risk-taking, but adding participants’ susceptibility to fear to the equation produced some interesting results. Although not significant, participants with a low sensitivity to fear had a tendency to fly more dangerously than those in the high sensitivity group. Interestingly, this tendency is swapped for those with more than 6 hours of weekly listening to music. If people are actually exposed to music for up to 28 hours a week (Webster, 2014), further research on exposure groups beyond 6 hours can be of great value.

Music preference could by itself not be related to participant risk-taking any more than weekly exposure. When split into fear susceptibility groups, however, a peculiar trait of these preferences came to light: Electronic music lovers showed a large disjunction in risk-taking, depending on their fear susceptibility. The high susceptibility group showed more speed-based risk-taking, and less distance-based risk-taking than the low susceptibility group, meaning that Electronic music lovers highly susceptible to fear actually flew more dangerously on two out of three risk measures. Although more distance was kept as well, this difference was smaller and less contrasting to the other data points.

The findings from this research add to the existing body of research around the Risk Homeostasis Theory in multiple ways. The evidence supporting RHT is strengthened and the controversy around the topic can be lessened. Aside from the topic itself, pathways to research combining it with musical cognition and emotional interference are opened up further, paving the way for a more complex understanding of the phenomenon. With a better understanding of risk homeostasis, a more accurate approach to new safety measures can be formed. Knowledge about the interpretation of these measures and possible emotional influence on its conversion to behavior could help us predict individuals’ reactions to danger and provide preventive instruction and/or measures. For example, different ways of displaying driving speed by enlarging the ‘red area’ on a speedometer.

Recreating the effect in a game environment also shows that risk homeostasis can occur in situations other than traffic. As risk-taking is an inseparable part of daily life, topics such as dangerous occupations, fire safety and stock trading are worth a scientific look.

4.5 Limitations

Despite interesting findings, some were defined by the limitations of this study. As described by Hoyes and Glendon (1993), an issue with the Risk Homeostasis Theory is how it's difficult to falsify. In the current situation, there is no way of learning how the target level of risk is determined for each individual and how measure it.

The experiment setting left some things to be desired as well. Although the controlled setting accounted for many factors, it inherently contains the issue of absence of actual harm, as raised by Hoyes & Glendon (1993). Because of this, participants may react to situations recreated in the experiment differently than they would have in the face of the actual risk. Another risk of a controlled experiment is participants exhibiting socially desirable behavior, which can cause less representative data.

Another limitation of this study was the duration of the experiment. With a practice round and five levels of up to four minutes each to play through, boredom will eventually get the better of participants. Due to the game's low complexity and singular goal ("do not get hit"), this is only a matter of time. How much time exactly is difficult to determine for each participant individually, but the effect manifests within sessions as well as throughout the experiment. This boredom can cause participants to make more risks than they usually would, as a way to end the game more quickly. Randomization of the conditions has been used to combat this on experiment-wide scale, but within sessions possible risky behavior due to boredom still remained.

Additionally, habituation time influenced the data at the beginning of each session. After starting, participants would need some time to decide the speed level that was comfortable for them. Because of this, speed is relatively low and stable in the 'first shield' part of the session, before a shield is lost for the first time. Furthermore, TTC and DCM contained irregularities in the beginning of sessions as well, as meteors have not yet started to fly in at that time. Because the two parameters are calculated using the distance of those meteors, the 'first shield' data is unrepresentative of actual risk-taking and was not included in the analyses. For speed, this data was included due to participants having control, but results may have been even more reliable if it were removed as well.

The biggest challenge to this research and its results was establishing a measure for the music-related variables. For hours of exposure, an approximation of expected listening was used, but results indicated that a wealth of information could have been hiding beyond '6 or more hours', had more separate categories been specified. For preference, the most commonly named genres were used for the experiment categories, but an almost infinite amount of various styles can be found upon searching. Grouping these together into collective styles might not only incur avid music lovers' wrath, but could also lead to data loss or misrepresentation, if not done carefully. Another issue with the music variables was the participants' level of involvement with it. Because most cognitive benefits of music have been shown to involve actual practice (Corrigall et al., 2013), not differentiating 'passive users' from 'players' and 'listening fanatics' can be a possible cause of inconclusive results.

The final issue, one that returned throughout the experiment, was the sample. Because it was small with only 69 participants, analyses had little power, significant results were less likely, and subgroups for music exposure and preference were unequally divided, with some containing as little as five cases. Furthermore, the sample contained little diversity. With the experiment performed at Leiden University's Faculty of Social Sciences, a considerable part of the participants consisted of Psychology students, mostly young females. A larger and more diverse sample would undoubtedly lead to better reliability and representativeness.

4.6 Future directions

Future research into the domain of Risk Homeostasis Theory can be of great value to further lessen the amount of controversy around it and contribute to the development of more sophisticated safety measures. To be able to learn more, current issues with the theory should be addressed.

Firstly, the falsification limitation, as raised by Hoyes and Glendon (1993). To be able to research whether the process taking place is actually risk homeostasis, a baseline of desired risk needs to be established. This baseline can differ for each individual, and thus needs to be determined as such. An example in the game environment could be: a number of short flying sessions without indication of protection, to determine a person's natural preference, repeated like a fingerprint registration to ensure a truthful window of risk.

As discussed earlier, the game environment only is not enough to be able to simulate realistic harm and thus produce generalizable results. To broaden the experimental risk-taking environment, a logical next step would be an experimental setting designed as close to real life as possible. In a traffic safety context, a driving simulator offers participants realistic and relatable input, and enables them to detect risk-taking feedback in a more natural manner, for example through a gas pedal or 'distance to obstacles' sensor in the virtual dashboard. For a more extensive range of risk-taking situations, virtual reality technology can offer a viable solution. Almost any environment can be simulated around participants, given the availability of a suitable medium for risk-taking measurement.

Continuing on from a measurable baseline of desired risk and a realistic environment, the amount of protection offered to participants must be noticeably different. In this experiment, risk homeostasis mainly occurred between levels of protection that were further apart, on a between-session scale. This could point to the levels of protection in the game not being distinguishable enough to be interpreted as less or more, resulting in minimal compensation. Including more distinct levels of protection to future experiments could result in a more pronounced homeostasis effect.

For research in the game environment, habituation time and boredom needs to be addressed as well. As realistic situations contain more factors than a game as well as a meaningful purpose (for example driving to arrive at work on time), taking risks is more a result of external factors and personal goals and less of simply being in these situations. Additionally, real-life tasks such as driving or performing a job usually don't need time to get used to, as opposed to the game controls. The issue of habituation can be tackled by disregarding the first few seconds of data, or creating an unmeasured 'ready, set, go' period at the start of a session altogether. The issue of boredom is a logical result of a controlled experimental setting with little factors, but difficult to address. Shortening the sessions would relieve some of the issue, but is a limited option as a certain bottom line of data gathering is necessary to determine reliable parameter means. Another way of reducing 'risk by boredom' would be to engage the player by including more meaningful and noticeable goals and feedback to the test environment. This could take the shape of a visually presented and recognizable goal (the flag to reach, or the person to help), a progress bar or the background changing when having flown a larger distance.

Future research on music as a factor of risk-taking offers great potential if performed carefully. As it is used intensively in daily life, it can be a valuable channel into safe behavior. For a more representative view on preference and exposure, a larger sample is necessary. A large sample produces more reliable results and will allow for the analysis of more, better filled preference and exposure groups. Results from a study including several preference groups provides specific insight into a larger part of the population and their risk-taking tendencies. For exposure, differentiating groups beyond 6 hours of weekly listening is expected to reveal valuable information on cases that were grouped together up until now. Adding music practice to exposure and preference as a factor could lead to a more in-depth understanding of the relation between music and behavior.

Incorporating the learnings from this study into future research will help us gain a better understanding of risk-taking and homeostasis, only to learn again from future attempts. This journey of constantly improving what we know and how we implement this knowledge will eventually bring us closer to a world where we can confidently be: safe and sound.

5. References

Adams, J. G. (1988). Risk homeostasis and the purpose of safety regulation. *Ergonomics, 31*(4), 407-428.

Apter, M. (1984). Reversal theory and personality: A review. *Journal of Research in Personality, 18*(3), 265-288.

Arnett, J. (1991). Heavy metal music and reckless behavior among adolescents. *Journal of Youth and Adolescence, 20*(6), 573-592.

Arnett, Jeffrey. (1992). The Soundtrack of Recklessness: Musical Preferences among Adolescents. *Journal of Adolescent Research, 7*(3), 313-31.

Aschenbrenner, M., & Biehl, B. (1985). Improved safety through improved technical measures? Empirical studies regarding risk compensation processes in relation to anti-lock brake systems. In R. M. Trimpop and G. J. S. Wilde (Eds), *Changes in accident prevention: The issue of risk compensation.* Groningen, The Netherlands: Styx publications.

Assum, T., Bjørnskau, T., Fosser, S., & Sagberg, F. (1999). Risk compensation—the case of road lighting. *Accident Analysis and Prevention, 31*(5), 545-553.

Ayres, T. J., & Hughes, P. (1986). Visual acuity with noise and music at 107 dBA. *The Journal of Auditory Research, 26*, 65-74.

Baniela, S. I., & Ríos, J. V. (2010). The Risk Homeostasis Theory. *Journal of Navigation, 63*(04), 607-626.

Brodsky, W. (2001). The effects of music tempo on simulated driving performance and vehicular control. *Transportation Research Part F: Psychology and Behaviour, 4*(4), 219-241.

Corrigall, K. A., Schellenberg, E. G. & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology, 4*, 1-10

Dezecache,G., Jacob, P., & Grèzes, J. (2015). Emotional contagion: Its scope and limits. *Trends in Cognitive Sciences, 19*(6), 297-299.

Discogs - Database and Marketplace for Music on Vinyl, CD, Cassette and More. Retrieved from https://www.discogs.com/ on February 22, 2016.

Doherty, R. W. (1997). The Emotional contagion scale: A measure of individual differences. *Journal of Nonverbal Behavior, 21*, 131-154.

Elvik, R. (2004). To what extent can theory account for the findings of road safety evaluation studies? *Accident Analysis and Prevention*, *36*(5), 841-849.

Enström, R. & Schmaltz, R. (2017). A Walk on the Wild Side: The Impact of Music on Risk-Taking Likelihood. *Frontiers in Psychology, 8*.

Evans, L. (1986). Risk homeostasis theory and traffic accident data. *Risk Analysis, 6*(1), 81-94.

Evans L. (1991). *Traffic safety and the driver*. New York, NY: Van Nostrand Reinhold.

Feggetter, A. J. (1982). A method for investigating human factor aspects of aircraft accidents and incidents. *Ergonomics, 25*, 1065-1075.

Fischhoff, B., Furby, L., & Gregory, R. (1987). Evaluating voluntary risks of injury. *Accident Analysis and Prevention, 19*(5), 1-62.

Forgas, J. P. (1998). On being happy and mistaken: Mood effects on the fundamental attribution error. *Journal of Personality and Social Psychology, 75*, 318–331.

Freud, S. (1975). *The psychopathology of everyday life* (A. Tyson, Trans). Harmondsworth, Penguin. (Original work published 1904).

Glendon, A. I., Hoyes, T. W., Haigney, D. E., & Taylor, R. G. (1996). A review of risk homeostasis theory in simulated environments. *Safety science, 22*(1), 15-25.

Grant, B.A., & Smiley, A. (1993). Driver response to anti lock brakes: A demonstration of behavioral adaptation. *Proceedings, Canadian Multidisciplinary Road Safety Conference VIII*, Saskatoon, Saskatchewan, June 14-16. Saskatoon, Saskatchewan, Canada: University of Saskatchewan.

Gravell, D. (2011, January 25). The fundamental music genre list. Retrieved from https://www.blisshq.com/music-library-management-blog/2011/01/25/fundamental-music-genre-list/

Gregersen,N. & Berg, H. (1994). Lifestyle and accidents among young drivers. *Accident Analysis and Prevention, 26*(3), 297-303.

Haight, F. (1986). Risk, especially risk of traffic accident. Accident *Analysis and Prevention, 18(*5), 359-366.

Hale, A. R. & Glendon, A. I. (1987). *Individual behaviour in the control of danger*. Amsterdam: Elsevier.

Halko, M., & Kaustia, M. (2015). Risk ON / Risk OFF: Risk-Taking Varies with Subjectively Preferred and Disliked Music. *PLoS ONE 10*(8): e0135436.

Halko, M., Mäkelä, T., Nummenmaa, L., Hlushchuk, Y., Schürmann, M., Houser, Daniel, & Weber, Bernd. (2015). Hedonic Context Modulates Risky Choices and Reward Responses in Amygdala and Dorsal Striatum. *Journal of Neuroscience, Psychology, and Economics*, *8*(2), 100-115.

Hatfield, E., Cacioppo, J., & Rapson, R. (1993). *Emotional Contagion (Studies in Emotion and Social Interaction)*. Cambridge: Cambridge University Press.

Horvath, P., & Zuckerman, M. (1993). Sensation seeking, risk appraisal, and risky behavior. *Personality and individual differences*, *14*(1), 41-52.

Hoyes, T. (1994). Risk homeostasis theory — Beyond transportational research. *Safety Science, 17*(2), 77-89.

Hoyes, T. W., & Glendon, A. I. (1993). Risk homeostasis: issues for future research. *Safety science*, *16*(1), 19-33.

Hoyes, T. W., Dorn, L., Desmond, P. A., & Taylor, R. (1996). Risk homeostasis theory, utility and accident loss in a simulated driving task. *Safety science*, *22*(1), 49-62.

Hoyes, T.W., Stanton, N.A., & Taylor, R.G. (1996). Risk homeostasis theory: A Study Of Intrinsic Compensation. *Safety Science, 22*(1), 77-86.

Isen, & Patrick. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance, 31*(2), 194-202.

Jackson, J. S., & Blackman, R. (1994). A driving-simulator test of Wilde's risk homeostasis theory. *Journal of Applied Psychology, 79*(6), 950.

Janssen, W., & Tenkink, E. (1988). Risk homeostasis theory and its critics: time for an agreement. *Ergonomics, 31*, 429-433.

Johnson, E., Tversky, A., & Manis, Melvin. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology, 45*(1), 20-31.

Josephs, R. A., Larrick, R. P., Steele, C. M., & Nisbett, R. E. (1993). Protecting the self from the negative consequences of risky decisions. *Journal of Personality and Social Psychology, 62*, 26-37.

Juslin, P. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews, 10*(3), 235-266.

Kahneman, Daniel. (2003). A perspective on judgment and choice: Mapping bounded rationality. (Author Abstract). *The American Psychologist*, *58*(9), 697-720.

Kreutz, G., Ott, U., Teichmann, D., Osawa, P., & Vaitl, D. (2008). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of Music, 36*(1), 101-126.

Lu, M. (2006). Modelling the effects of road traffic safety measures. *Accident Analysis and Prevention, 38 (3)*, 507–517.

Maughan-Brown, B., & Venkataramani, A. (2012). Learning that circumcision is protective against HIV: Risk compensation among men and women in Cape Town, South Africa. *PloS One*, *7*(7), E40753.

Mckenna, F. P. (1987). Behavioural compensation and safety. *Journal of Occupational Accidents*, *9*(2), 107-121.

Miller, J. Gregg, Jr. (1998). Risk homeostasis and California design defect products liability: Rethinking the consumer expectations and risk-benefit tests. *University of San Francisco Law Review, 32*(3), 587-627.

O'Neill, B., & Williams, A. (1998). Risk homeostasis hypothesis: a rebuttal. *Injury Prevention, 4*(2), 92-93.

Sawyer, J. E., Kernan, M. C., Conlon, D. E., & Garland, H. (1999). Response to the Michelangelo computer virus threat: The role of information sources and risk homeostasis theory. *Journal of Applied Psychology, 29*(1), 23-51.

Schulze, H. (1990) Lifestyle, leisurestyle and traffic behaviour of young drivers. In: *VT1 report 364A.* Linkaping: Swedish Road and Traffic Research Institute (VTI).

Shannen, H. S., & Szatmari, P. (1994). Seat-belt legislation and risk homeostasis: Further analysis of the British Data. *Accident Analysis & Prevention, 26*(6), 803-805.

Simonet, S., & Wilde, G. J. (1997). Risk: Perception, acceptance and homeostasis. *Applied Psychology, 46*(3), 235-252.

Spinney, L. (1997) Pump down the volume. *New Scientist*, *155*, 22

Stalder, D., & Cook, J. (2014). On Being Happy and Mistaken on a Good Day: Revisiting Forgas' (1998) Mood-Bias Result. The Journal of Social Psychology, *The Journal of Social Psychology, 137*(1), 371-374.

Stetzer, A., & Hofmann, D. (1996). Risk compensation: Implications for safety innovations. *Organizational Behavior and Human Decision Process, 66*(1), 73-88.

Swuste, P., Groeneweg, J., Van Gulijk, C., Zwaard, W., & Lemkowitz, S. (2017). Safety management systems from Three Mile Island to Piper Alpha, a review in English and Dutch literature for the period 1979 to 1988. *Safety Science,* Safety Science.

The most comprehensive list of genres of music available on the Internet. (2015). Retrieved from http://www.musicgenreslist.com/

Trimpop, R. (1996). Risk homeostasis theory: Problems of the past and promises for the future. *Safety Science, 22*(1), 119-130.

Webster, T. (2014). Edison Research Conducts First Ever Share of Ear® Measurement For All Forms Of Online And Offline Audio [blog post]. Retrieved from: http://www.edisonresearch.com/edison-research-conducts-first-ever-share-of-ear-measurement-for-all-forms-of-online-and-offline-audio/

Wilde, G. J. (1989). Accident countermeasures and behavioural compensation: The position of risk homeostasis theory. *Journal of Occupational Accidents*, *10*(4), 267-292.

Wilde, G. J. (1982). The theory of risk homoeostasis: implications for safety and health. *Risk analysis, 2*(4), 209-225.

Wilde, G. J. S. (1998). Risk homeostasis theory: an overview. *Injury Prevention, 4*(2), 89–91.

Wilde, G. J., Robertson, L., & Pless, I. B. (2002). Does risk homoeostasis theory have implications for road safety. *BMJ (Clinical Research Ed.), 324*(7346), 1149-1152.

Yuen, K. S. L., & Lee, T. M. C. (2003). Could mood state affect risk-taking decisions? *Journal of Affective Disorders, 75*(1), 11-18.

Velden, van der L. (2017, June 30). Douwe Egberts staakt reclamecampagne met jongeren die van brug springen. *De Volkskrant*. Retrieved from https://www.volkskrant.nl/

Vinagre-Rios, J., & Iglesias-Baniela, S. (2013). The Human Element in Shipping Casualties as a Process of Risk Homeostasis of the Shipping Business. *Journal of Navigation, 66*(3), 385-397.

Appendix A: Questionnaire Music Preference

1. How many hours per week do you listen to music?

- ☐ 0 - 0,5 hours
- ☐ 0,5 - 2 hours
- ☐ 2 - 4 hours
- ☐ 4 - 6 hours
- ☐ >6 hours

2. From this list of music genres, which 3 would you say you prefer most?

- ☐ Classical
- ☐ Electronic
- ☐ Funk
- ☐ Soul
- ☐ Hip-hop
- ☐ Jazz
- ☐ Latin
- ☐ Pop
- ☐ R&B
- ☐ Reggae
- ☐ Rock
- ☐ Rap
- ☐ World
- ☐ Other (please specify) …

3. From the 3 genres you have chosen, please construct a top 3.

1) […] 2) […] 3) […]

Appendix B: Doherty's (1997) Emotional Contagion Scale, as presented in the questionnaire

The next questions will be about various of your feelings and behaviors in various situations. There are no right or wrong answers, so try very hard to be completely honest in your answers. Results are *completely confidential.* Read each question and indicate the answer which best applies to you. Please answer each question very carefully. Thank you.

Use the following key:

4. *Always* = Always true for me.
3. *Often* = Often true for me.
2. *Rarely* = Rarely true for me.
1. *Never* = Never true for me.

1. If someone I'm talking with begins to cry, I get teary-eyed.
Never 1 2 3 4 Always

2. Being with a happy person picks me up when I'm feeling down.
Never 1 2 3 4 Always

3. When someone smiles warmly at me, I smile back and feel warm inside.
Never 1 2 3 4 Always

4. I get filled with sorrow when people talk about the death of their loved ones.
Never 1 2 3 4 Always

5. I clench my jaws and my shoulders get tight when I see the angry faces on the news.
Never 1 2 3 4 Always

6. When I look into the eyes of the one I love, my mind is filled with thoughts of romance.
Never 1 2 3 4 Always

7. It irritates me to be around angry people.
Never 1 2 3 4 Always

8. Watching the fearful faces of victims on the news makes me try to imagine how they might be feeling.
Never 1 2 3 4 Always

9. I melt when the one I love holds me close.
Never 1 2 3 4 Always

10. I tense when overhearing an angry quarrel.
Never 1 2 3 4 Always

11. Being around happy people fills my mind with happy thoughts.
Never 1 2 3 4 Always

12. I sense my body responding when the one I love touches me.
Never 1 2 3 4 Always

13. I notice myself getting tense when I'm around people who are stressed out.
Never *1* *2* *3* *4* *Always*

14. I cry at sad movies.
Never *1* *2* *3* *4* *Always*

15. Listening to the shrill screams of a terrified child in a dentist's waiting room makes me feel nervous.
Never *1* *2* *3* *4* *Always*

Appendix C: Calculating the risk parameters

Three risk-taking parameters were calculated from the game data: flying speed in pixels per second (speed), the time to collision in seconds with the first meteor in its path (TTC), and the distance in pixels to the closest meteor anywhere on the screen (DCM).

These were calculated using variables the game collected in the 'steplog.csv' file.

For calculating speed in pixels, the variable 'difficulty' was used. Difficulty represented the speed level from 1 (320 pixels per second) to 13 (920 pixels per second), with an increase of 50 pixels per second per step. The formula used was as follows:

$$speed_p = \left(2.7 + (difficulty * 0.5)\right) * 100$$

For calculating TTC, the variables 'meter in path location x' and 'speed' were used. Respectively, these are the distance to the closest meteor on the spaceship's path (in pixels from the left border) and the calculated speed in pixels per second. The 109 pixels that are subtracted before dividing represent the horizontal position of the spaceship from the left border, as its position is slightly to the right of it. The following formula was used:

$$TTC = \frac{meteor\ in\ path\ location\ x - 109}{speed_p}$$

For calculating DCM, the variables closest meteor location x', 'closest meteor location y' and 'ship location y' were used. Respectively, these are the closest meteor horizontally (in pixels from the left border), the closest meteor vertically (in pixels from the top border), and the ship location in pixels from the top border.
The formula used to compute ‘distance to the closest meteor’ (DCM) was derived from the Pythagorean Theorem and goes as follows:

$$distance = \sqrt{(closest\ meteor\ location\ x)^2 + (closest\ meteor\ location\ y - ship\ location\ y)^2}$$

Appendix D: The instructions letter

Information letter - Risk homeostasis in gaming

Welcome and thank you for coming! You are going to play a computer game and fill in a questionnaire. Before you start, please read this information letter and sign the informed consent. Your participation is completely anonymous and voluntarily. Your records are coded by means of a participant number (see the post-it). You will need to enter this number when starting the game and the questionnaire. Please double check when entering your number, this is important. If you would like to stop the experiment you may do so at any moment. The results of this study will be used in SPSS to conduct statistical analyses for our master thesis about the risk homeostasis theory.

The game

The game is about a little spaceship in a galaxy not so far away on its way to deliver very valuable cargo. The spaceship is in a hurry and has to reach its destination as soon as possible. Unfortunately, the ship runs into a thick cloud of meteors. You are the ship's captain and you have to stay on your toes to dodge the danger and get through. The goal is to go as fast as you can (a faster speed will result in more points) but also try to avoid the meteors (a collision with a meteor will cost you a life).
You will receive specific instructions about the game (e.g. which buttons to use etcetera) when starting the game.

Instructions

Please pay attention only to your own computer screen. Also, please do not make noise. When you have a question raise your hand and one of us will come to you.

After you have read and completed the informed consent, please login with your UL account (Some of the computers are already logged in, if so, do not log in with your own UL account). When your desktop is completely loaded raise your hand. We will start the game for you After you have finished the game please raise your hand and we will start the questionnaire for you. *Please do not forget to enter your (correct) post-it number both in the game and questionnaire*! When you completed the questionnaire you can collect your money or credits for participating.

Any questions?

Remarks or complaints afterwards can be directed towards the senior researcher:

Jop Groeneweg
Groeneweg@fsw.leidenuniv.nl

Appendix E: The informed consent form

Informed Consent - Risk homeostasis in gaming
In this experiment we will test the risk homeostasis theory by means of a computer game. The experiment will take about 45 minutes. You will be compensated for your time by receiving 2 credits or €6,50. By signing the form you agree with the following statements.

- I have read the information letter. I could ask additional questions. Questions that I had have been answered adequately. I have had sufficient time to decide whether or not I participate.

- I am aware that participation is completely voluntary. I know that I can decide at any moment not to participate or to stop. I do not need to provide a reason for that.

- My responses are processed anonymously or in a coded way.

- I give consent to use my data for the purposes that are mentioned in the information letter.

I consent to participating in this study.

Name of participant: ____

Signature: ____

Date: _____/_____/______