Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Risk Homeostasis Theory and the Effect of Affect on Risk-Taking and Performance

Celina Mons

Master's Thesis Applied Cognitive Psychology
Faculty of Social and Behavioral Sciences, Leiden University

Student number: s1109758
E-mail address: c.mons@live.nl
Supervisor: Dr. Jop Groeneweg
September 2017

Table of Contents

Abstract

Since its publication in 1982, Wilde's Risk Homeostasis Theory (RHT) has been the subject of controversy amongst scientists. The current study attempted to test this much-debated theory in a laboratory experiment, using a video game in which participants could freely adjust the amount of risk they were taking and received varying amounts of protection. Additionally, the effect of affect on RHT was investigated by measuring the effect on risk-taking and performance. The data showed no evidence for or against the theory, however, they highlighted the difficulties of testing RHT in general. Affect seemed to have an influence on performance, unfortunately, the actual impact on RHT remains unclear due to the difficulties of testing the theory. Suggestions to improve the current experiment and to test RHT in the future are discussed.

Keywords: risk homeostasis, risk compensation, risk-taking, affect, performance, attention

1. Introduction

Risks are a part of human life. Every day millions of people travel on busy streets, smoke cigarettes, gamble, migrate, and so on. Aside from these risks that people are voluntarily taking, there are the risks of, for example, terrorist attacks and natural disasters. Luckily, humans are quite inventive when it comes to developing safety measures that decrease the amount and magnitude of harmful outcomes. Right?

Wilde (1982) published an article in which he argued that added safety measures lead people to adjust their behavior in a way that the added safety is compensated for. More specifically, his *Risk Homeostasis Theory* (RHT) states that there is a certain level of risk that people are willing to take: the target level of risk. In addition, there is a perceived level of risk. People want to balance both levels, but if a safety measure is added the perceived level of risk decreases, yet the target level of risk remains constant. To restore the balance, people will eventually alter their behavior until the perceived level of risk again matches the target level of risk. Thus, the number of harmful outcomes will not decrease, unless one's target level of risk decreases. This level depends on motivation and the perceived costs and benefits of certain behavior, but is not altered by the addition of safety measures. RHT was initially developed to explain traffic related behavior, however, according to Wilde it is also applicable to other behavioral domains that hold health or safety implications. A model of the homeostatic mechanism, reprinted from Wilde (1998, p. 90), is displayed in Figure 1.

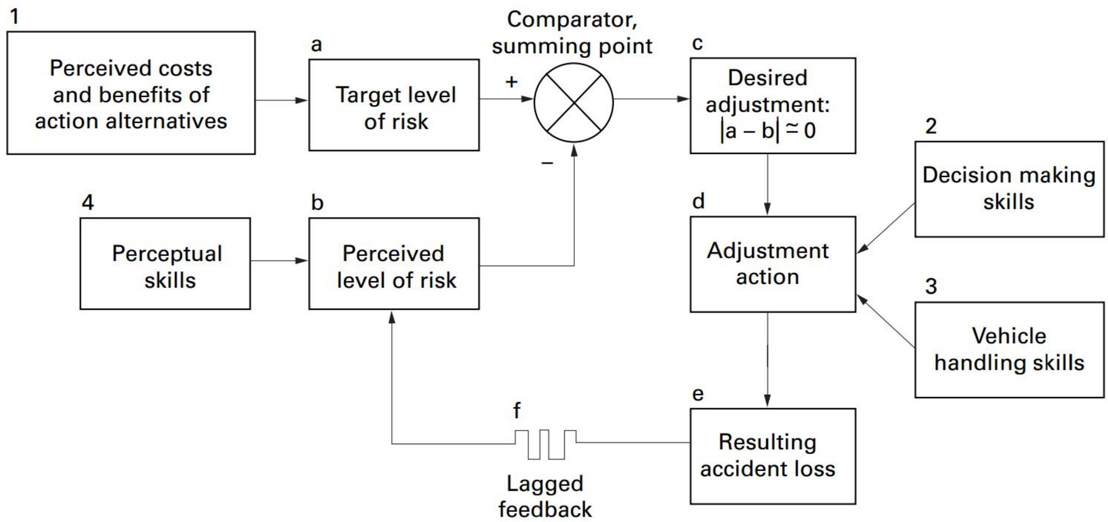

Figure 1. Homeostatic mechanism. Reprinted from "Risk homeostasis theory: an overview," by G. J. S. Wilde, 1998, *Injury Prevention, 4*(2), p. 90, Copyright 1998 by BMJ Publishing Group.

This model of the homeostatic mechanism (Figure 1) will be used in the following example of RHT: When people are riding a bicycle, they are willing to take a certain amount of risk (*a*) of having an accident and sustaining a head injury (*e*). If they start wearing a helmet, cycling becomes less risky (*b*). At first, the imbalance between *a* and *b* might go unnoticed and lead to fewer head injuries (*e*). However, once the imbalance is noticed, people will try to eliminate it (*c*) by riding their bicycles in a more dangerous fashion (*d*). This change in behavior will eventually cancel out any added safety provided by the helmet. Therefore, the number of head injuries will rise back to its original level (*e*). Only if people reevaluate the costs and benefits of their behavior, through which their motivation and target level of risk change (*1* altering *a*), there will be a permanent change in accident loss (*e*).

1.1. Debating Risk Homeostasis Theory

Throughout the years, an extensive amount of research has been conducted regarding RHT. Interestingly, scientists cannot agree on whether to accept or to reject the theory. Wilde (1982) found support for his theory in existing data on traffic accidents. For example, RHT can explain why newly implemented safety regulations, such as improved vehicle lightning, the installation of seatbelts and three phase traffic lights, ultimately had no effect on accident rates. Furthermore, an experiment with a driving simulator, conducted by Jackson and Blackman (1994), supports RHT. Their data showed that manipulating factors that influence motivation had an inverse effect on the number of accidents, while manipulating nonmotivational factors showed no effect. This finding supports the notion that one's target level of risk is essential to modulate accident rates. Additional support for RHT comes from Baniela and Ríos (2010), who analyzed maritime accident data from 2005 and 2006. They found that ships, regardless of their objective safety, suffered the same amount of accidents. Therefore, it was concluded that safety measures did not alter the crew's target level of risk, resulting in behavioral compensation for the added safety.

In contrast, major criticism on RHT was expressed by Evans (1986), who analyzed a variety of traffic data showing either no support for the theory or contradicting it. Particularly striking is his finding that some data that were previously thought to support RHT, have shown to be inconsistent with the theory. Evans therefore advises to reject RHT. O'Neill and Williams (1998) agree. They mainly criticize RHT for its proposed feedback mechanism, which they think is too complex to exist. In an experiment by Hoyes, Dorn, Desmond, and Taylor (1996) RHT was also refuted. The data obtained in their driving simulator study opposed the theory's notion that a change in behavior does only occur when costs or benefits are perceived.

Adams (1988), on the other hand, doubts the general falsifiability of RHT. Although he finds the theory plausible, he thinks of it as a metaphysical concept and therefore impossible to falsify. Hoyes and Glendon (1993) share this skepticism. They state that, due to the vague definition of accident loss and motivational change, it is impossible to falsify RHT. The theory must be formulated in a way that makes it possible to objectively support or disprove it. Wilde (2014) states that there are indeed some difficulties when it comes to testing RHT, however, those difficulties are not due to its formulation. They rather arise because of ethical restrictions when it comes to testing RHT in real-life situations. Therefore, one is forced to rely on existing traffic data and, although suboptimal, laboratory research.

It is evident that, since its publication, RHT has sparked controversy. Considering that there is still no unified opinion about its validity, RHT remains an interesting topic for investigation. To test RHT in the present study, a video game will be programmed in which participants will be able to freely adjust the amount of risk they are taking throughout the game, and in which they will be given varying amounts of protection. Additionally, assuming that people are taking more risk in order to maximize their benefits, it will be investigated if taking more risk is truly beneficial for reaching one's goal. In other words, if taking a higher risk is related to higher game performance.

1.2. The role of affect in RHT

Affect has shown to be a force to be reckoned with. Various studies demonstrated its effect on, for example, decision making (Finucane, Alhakami, Slovic, & Johnson, 2000), risk-taking (Isen and Patrick, 1983), the attentional scope (Fredrickson & Branigan, 2005), and even auditory perception (Siegel & Sefanucci, 2011). Therefore, the role of affect in RHT will be explored in the current study, specifically its effect on the two levels of risk. Figure 1, displaying the homeostatic mechanism, will be used to illustrate the reasoning in the following two paragraphs.

1.2.1. Establishing a target level of risk. One very important part of Wilde's (1982) RHT is the target level of risk (*a*), which is proposed to be essential to a permanent change in accident loss (*e*), and is based on the perceived costs and benefits of one's behavior (*1*). For example, when it comes to driving speed, one might attach more value to avoiding a speeding ticket than to gaining time by driving faster and vice versa. Although this seems quite rational and straightforward, people also tend to rely on their affective feelings when making judgments about costs and benefits. Finucane et al. (2000) proposed the *affect heuristic,* meaning that people use their affective feelings towards certain situations, objects, etcetera as a decision-making shortcut. To demonstrate the use of this heuristic they performed two experiments. In the first experiment, participants were asked to make judgments about the risks and benefits of several activities and technologies. It was hypothesized that the often-observed negative correlations between subjective risks and benefits ratings would be found, and that they would be stronger when participants were under time pressure. The data supported the hypothesis by showing that people rely on the affect heuristic and that this effect gets stronger when there is less time for analytic judgment. In their second experiment, Finucane et al. (2000) asked participants once more to make judgments about the risks and benefits of certain technologies. Then, the participants were informed about either the risks or the benefits of these technologies. The results again showed the inverse relationship between risks and benefits. But they also showed that when one of the affective attributes was successfully manipulated, the other would change in the opposite direction. For example, if people judged the risks of nuclear power plants as high and the benefits as low, but were later convinced that nuclear power plants have great benefits, their rating of the risks decreased. Throughout the years, the affect heuristic has received additional support (Slovic, Peters, Finucane, & MacGregor, 2005; Slovic, 2010).

Furthermore, Isen and Patrick (1983) conducted a gambling experiment in which they informed participants about the risk of losing a bet. They found that participants in a positive affective state, compared to participants in a neutral affective state, placed higher bets when the risk of losing was low. But when the risk of losing was high, the positive affect group took less risk than the neutral affect group. However, in a simultaneously performed experiment Isen and Patrick (1983) let participants read hypothetical dilemmas with a low, medium or high risk. It was found that participants in a positive affective state were more willing to take a high risk. Since hypothetical dilemmas form no actual threat, the authors suggest that the results of the two studies are consistent in the sense that participants in the positive affect group were only willing to take a higher risk in situations that form low risks in real-life. A more recent study by Yuen and Lee (2003) showed similar results. When asked to make a choice in a hypothetical life dilemma, participants in a positive mood were more inclined to take high risks than participants in a negative mood.

It is evident that affective feelings have at least some effect on costs and benefits judgments and therefore on risk-taking. In RHT terms, the experiments by Isen and Patrick (1983) and Yuen and Lee (2003) showed that when people in different affective states were given the same explicit level of risk (the perceived level of risk (*b*)), their willingness to take certain risks differed. Thus, their target level of risk (*a*) has likely been altered by their affective state. Hence, it is argued that affect plays a role in establishing the target level of risk (*a*) in RHT by modulating one's perception of costs and benefits (*1*). If this is true, differences in affective states will lead to differences in risk-taking in the current experiment. Since playing a video game poses a low to no real-life threat, it is expected that a positive affective state is related to the perception of high benefits (*1*), thus a higher target level of risk (*a*). The imbalance between the two risk levels (*a* and *b*) leads to the desire to adjust one's behavior (*c*), thus increased risk-taking (*d*). In other words, participants in a high positive affective state are expected to take more risk than participants in a low positive affective state.

1.2.2. Affecting the perceived level of risk. In his 1982 article, Wilde briefly displays the influence of distractions, or the lack of attention for a task, on the perceived level of risk. If people are distracted their perceptual skills decrease (*4*), through which their perceived level of risk changes (*b*), which creates an imbalance between the perceived (*b*) and the target level of risk (*a*). As a result, the homeostatic mechanism kicks in and people adjust their behavior (*d*) in an effort restore the balance (*c*). Interestingly, research has shown that affect also has an influence on attention. According to the *Broaden-and-Build Theory of Positive Emotions* "positive emotions broaden the scope of attention and thought-action repertoires" (Fredrickson & Branigan, 2005, p. 1). This notion was supported by two experiments performed by Gasper and Clore (2002). In the first experiment participants reproduced ambiguous drawings, and in the second experiment they rated whether geometric figures bore more resemblance to a figure with similar global aspects or a figure with similar local aspects. Both experiments showed that participants in a positive affective state payed more attention to global features, while participants in a negative affective state payed more attention to local features. Other experiments also found that positive affect broadens the scope of attention (Gasper, 2004; Fredrickson & Branigan, 2005; Rowe, Hirsh, & Anderson, 2007). Nevertheless, Rowe et al. (2007) point out that the broadened scope of attention also leads participants to being more easily distracted by irrelevant stimuli.

In the light of RHT, a broadened scope of attention can result in two things. First, it can lead to people having a broader overview of the situation. For example, if people are driving in a car, this can be beneficial because they might be better at recognizing dangerous situations. On the other hand, it can also result in people being more distracted by irrelevant stimuli, such as a big bright billboard. Vice versa, a narrower scope of attention might decrease one's susceptibility for distractions, but could be detrimental for overseeing the situation. Although it is not possible to measure the exact attentional scope and its influence on perceptual skills (*4*) and therefore the perceived level of risk (*b*), it is possible to explore its effect on performance. Since the current experiment will take place in a controlled environment, there should be very limited distractions. Therefore, a broader scope of attention is expected to be most beneficial because it leads to a broader overview of the game. That is why it is expected that participants in a high positive affective state will perform better on the current video game than participants in a low positive affective state.

1.3. Summarizing the aim of the present study

The main objective of the present study is to investigate Wilde's (1982) controversial Risk Homeostasis Theory. Therefore, a video game will be programmed in which participants will be able to freely adjust the amount of risk they are taking throughout the game, and in which they will be given varying amounts of protection. It is hypothesized that participants will behave as predicted by RHT (first hypothesis), thereby giving Wilde's theory the benefit of the doubt. Secondary, it hypothesized that taking a higher risk is related to higher game performance (second hypothesis). In addition, the effect of affect in RHT will be investigated, using the following two hypotheses: participants in a high positive affective state will take more risk than participants in a low positive affective state (third hypothesis), and, participants in a high positive affective state will perform better than participants in a low positive affective state (fourth hypothesis).

2. Method

2.1. Participants

The total number of participants in this study was 178. However, due to technical difficulties, the data from six participants could not be used in the analysis. Out of the remaining 172 participants 44 were male and 128 were female with ages ranging from 18 to 57 years (*M* = 22.42, *SD* = 4.72). Participants were recruited through SONA-systems and Facebook. For their participation in this 60-minutes experiment participants either received €6,50 or 2 course credits.

2.2. PANAS

The scale used to measure affect was the Positive and Negative Affect Schedule (PANAS) by Watson, Clark and Tellegen (1988). The PANAS consists of 20 words that can be used to describe one's feelings; half of the items measuring positive affect and the other half measuring negative affect. The items are rated on a scale ranging from 1 (very slightly or not at all) to 5 (extremely). In this study participants were asked to indicate how they were feeling at that given moment. The PANAS was presented to participants through Qualtrics.

2.3. Game

2.3.1. Gameplay. A video game, featuring a spaceship that had to be safely navigated through a meteor shower in order to deliver a valuable package, was used to assess risk-taking behavior. The game was created in GameMaker 8.1 and could be played on a computer by using the arrow keys on the keyboard. Figure 2 shows a screenshot of the game, including screen and object dimensions. Meteors entered the screen from the right and moved horizontally towards the ship on the left, but they could not move in the vertical direction. Pressing the right key increased the difficulty level of the game, which was linked to the speed at which the meteors moved across the screen, from level 1 (320 pixels per second) to level 13 (920 pixels per second). Pressing the left key decreased the difficulty level within the same range. Each round started with difficulty level 1, thus a meteor speed of 320 pixels per second. By pressing the up and down keys the ship could be moved in a vertical direction to avoid the incoming meteors. The ship's horizontal position was fixed. Although the meteors moved towards the ship, instead of the ship moving towards the meteors, it was made believable to the players that the ship was the only controlled object. Thus, players believed that they changed the speed of the ship instead of the meteor speed. This effect was accomplished by simultaneously moving the background and the meteors.

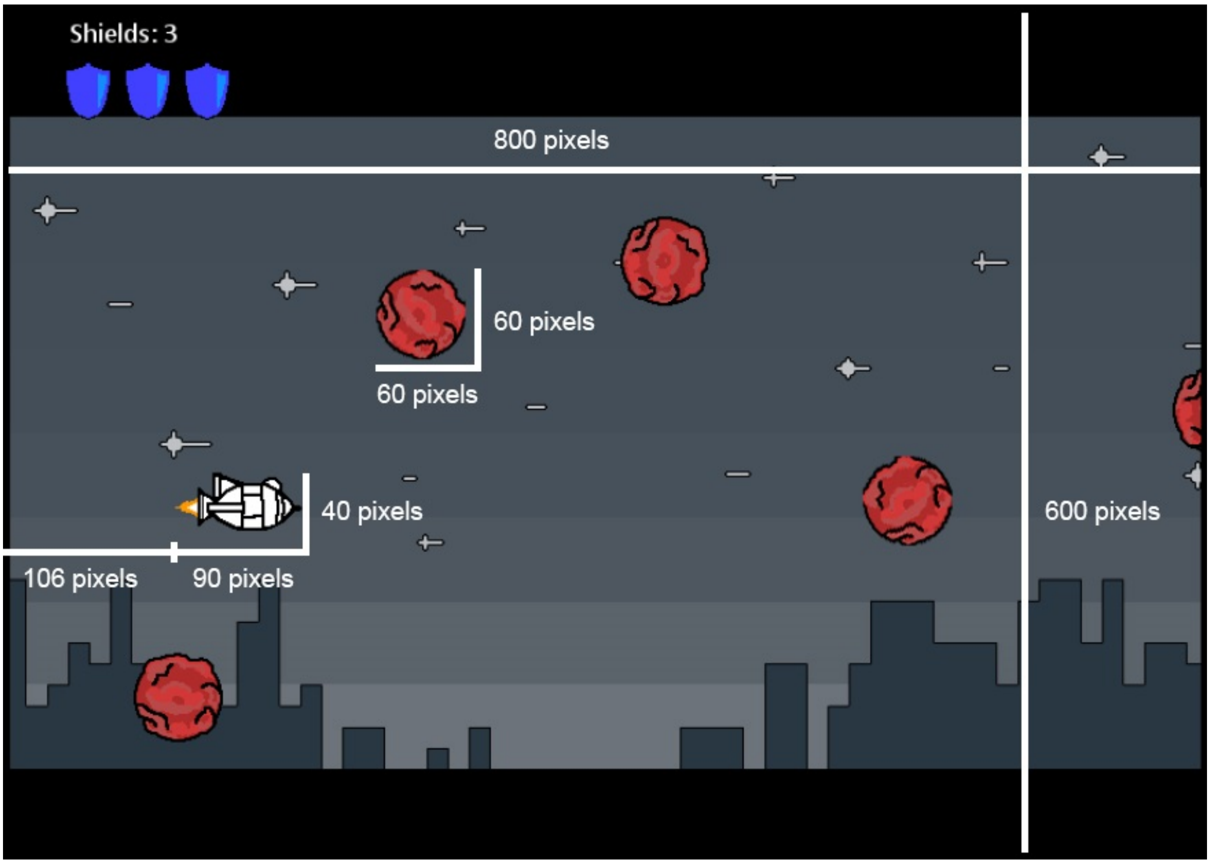

Figure 2. Screenshot of the game, including screen and object measurements.

The game consisted of five rounds, preceded by one practice round. After 4 minutes of playing, the ship would reach its destination and one round was completed. However, as an incentive to fly faster, participants were told that they had to cover a certain distance to finish the rounds. Rounds could also end prematurely. This happened when the ship exploded after colliding with a meteor. To prevent explosions, the ship was equipped with shields. Thus, when hitting a meteor while being equipped with shields, the ship would lose one of its shields and continue its journey. The number of shields left was displayed in the upper left corner of the screen and could start with 0, 1, 3, 5 or unknown to the player, depending on the condition. However, in the unknown number of shields condition, the ship was always equipped with three shields. All conditions were played once by every participant, but the sequence in which they were played was randomly allocated. During the practice round the number of shields was either one or three and was also randomly assigned.

2.3.2. Data storage and computation. During gameplay data was stored in two kinds of files to be used in the analysis: steplogs and eventlogs. Steplogs were files with data saved every 10 milliseconds and eventlogs were files with data saved whenever a collision occurred. Therefore, eventlogs showed the exact time and coordinates of the collision. The following data was stored in the files: the participant number, the condition, the number of shields, the difficulty level (linked to speed), the time (in seconds passed since the start of the game), the ship's location, the location of the closest meteor in path (the meteor that was currently in the ship's path and could cause a collision) and the location of the closest meteor (the meteor closest to the ship regardless of if it was in the ship's path). These data points were used for the computation of performance and risk-taking variables that were necessary for the analysis.

A game score could be calculated by multiplying the average speed with the time played during a session. This is also the distance travelled by a participant's spaceship. The score (or distance) was a measure of performance. Additionally, there were three main measures of risk-taking. The first one was speed, with higher speed indicating a higher risk of colliding with a meteor, thus more risk-taking. The second one was the *time to collision* (TTC). This was the time it would take to collide with the meteor in the ship's path if no action was taken. A lower time to collision indicated more risk-taking. TTC could be calculated by subtracting the horizontal location of the ship's bow (which was fixed at 196 pixels) from the horizontal location of the closest meteor in path (measured from the left side of the meteor), then dividing this number by speed. This resulted in the following formula used to calculate TTC:

$$TTC = \frac{location\ x\ closest\ meteor\ in\ path - 196}{speed}$$

The third main measure of risk-taking was how closely participants would navigate the ship past meteors, this was called the *distance to closest meteor* (DCM). Navigating the ship closer past meteors would indicate more risk-taking. To calculated the DCM the location of the ship and the location of the closest meteor were used in the following formula:

$$DCM = \sqrt{(closest\ meteor\ location\ x - 196)^2 + \ (closest\ meteor\ location\ y - ship\ location\ y)^2}$$

2.4. Procedure

Participants were tested in a computer lab at the Faculty of Social and Behavioral Sciences of Leiden University. They were seated behind a computer, with at least one empty seat between participants. Once they had read and signed the informed consent form, participants received further instructions about the experiment. They were told about the goal of the game (delivering a valuable package) and that they could earn more points when the package was delivered quickly. In order to motivate the participants to perform well on the task, they were informed that the participant with the highest overall score would receive an extra reward (€25,- or a lottery ticket of the same value). Then, the experiment started. First the participants were asked to fill out the PANAS along with a questionnaire asking for their age, gender and participant number (which they could find on a note on their desk). Then, the game started. Once again, participants were asked to fill in their participant number. After they had played the game, participants received their reward (€6,50 or 2 course credits), a snack and a debriefing form. Once the scores of all participants were compared, the participant with the highest score received the extra reward.

3. Results

Microsoft Excel was used to convert the steplogs and eventlogs into a file that could be used in SPSS. The data obtained through Qualtrics could be downloaded in SPSS format and was merged with the game data. All statistical analyses were performed in IBM SPSS Statistics 23. A significance level of $\alpha = .05$ was used.

3.1. First hypothesis: Participants will behave as predicted by RHT

3.1.1. Mean speed. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity ($\chi^2(135) = 1266.20$, $p < .001$). Therefore, a Repeated Measures ANOVA with Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the mean speed ($F(6.59, 757.49) = 49.91$, $p < .001$, $\eta_p^2 = .30$). However, the post-hoc Bonferroni correction showed that only the first shield of every round significantly differed from the rest of the non-starting shields (all *p*s < .001), with the unknown number of shields condition being an exception. In this condition, the second shield also significantly differed from the other shields in its round (all *p*s < .01). The other differences between shields were not significant (all *p*s > .05). This means that the mean speed was significantly lower for all first shields compared to any other shield in their condition, but that non-starting shields did not significantly differ on mean speed. With the second shield in the unknown number of shields condition being an exception. For this shield, the mean speed was significantly higher compared to the first shield, and significantly lower compared to the remaining shields of this round. A graphic representation of the number of shields in relation to mean speed is shown in Figure 3.

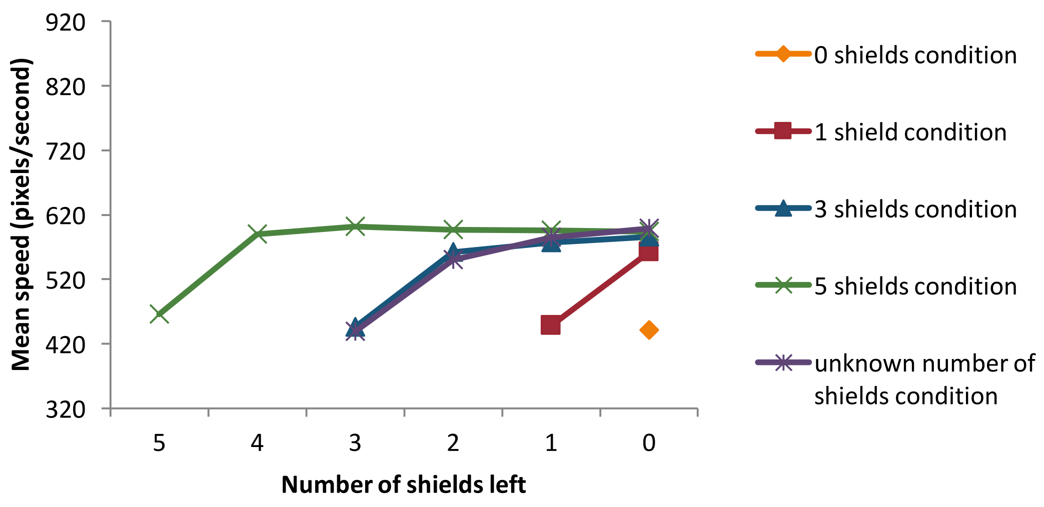

Figure 3. The relationship between the number of shields and the mean speed, with the vertical axis ranging from the lowest to the highest possible speed.

3.1.2. Maximum speed. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity ($\chi^2(135) = 1085.53$, $p < .001$). Therefore, a Repeated Measures ANOVA with Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the maximum speed ($F(7.47, 858.44) = 9.11$, $p < .001$, $\eta_p^2 = .07$). However, the post-hoc Bonferroni correction showed that in the one shield condition there was no significant difference between having or not having a shield ($p = .051$). In the three shields condition the only significant differences emerged between the first and the second shield ($p = .01$), and the first and the third shield ($p = .049$). The other differences in this round were not significant (all *p*s > .05). A similar result was found in the five shields condition, where the first shield also significantly differed from the second ($p = .03$) and the third shield ($p < .001$), but no other differences occurred (all *p*s > .05). In the unknown number of shields condition the first shield significantly differed from all other shields (all *p*s < .001), while no differences arose between the other shields in this round (all *p*s > .05). This means that the maximum speed for the first shield of the three and five shields conditions was significantly lower than for the second and third shields of those rounds. In the unknown number of shields condition the maximum speed measured during the first shield was significantly lower than all other maximum speeds measured in this round. The other shields did not significantly differ on maximum speed. A graphic representation of the number of shields in relation to maximum speed is shown in Figure 4.

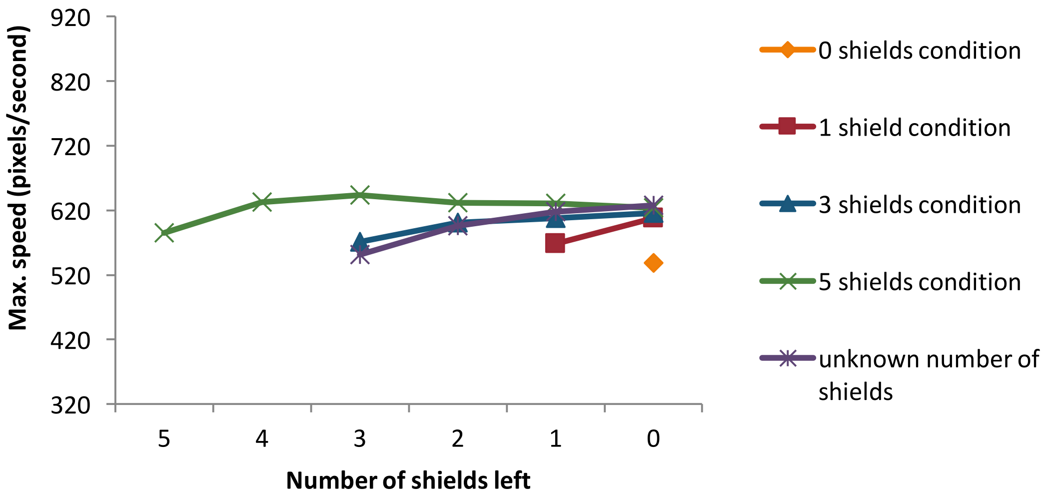

Figure 4. The relationship between the number of shields and the maximum speed, with the vertical axis ranging from the lowest to the highest possible speed.

3.1.3. Mean TTC. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity ($\chi^2(135) = 429.85$, $p < .001$). Therefore, a Repeated Measures ANOVA with

Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the mean TTC (F(10.46, 1202.65) = 75.52, p < .001, η_p^2 = .40). However, the post-hoc Bonferroni correction showed that, as before for mean speed, only the first shield of every round significantly differed from the rest of the non-starting shields (all ps < .001). No other significant differences were found (all ps > .05). This means that the mean TTC was significantly higher for all first shields compared to any other shield in their condition, but that the non-starting shields did not significantly differ on mean TTC differ. A graphic representation of the number of shields in relation to mean TTC is shown in Figure 5.

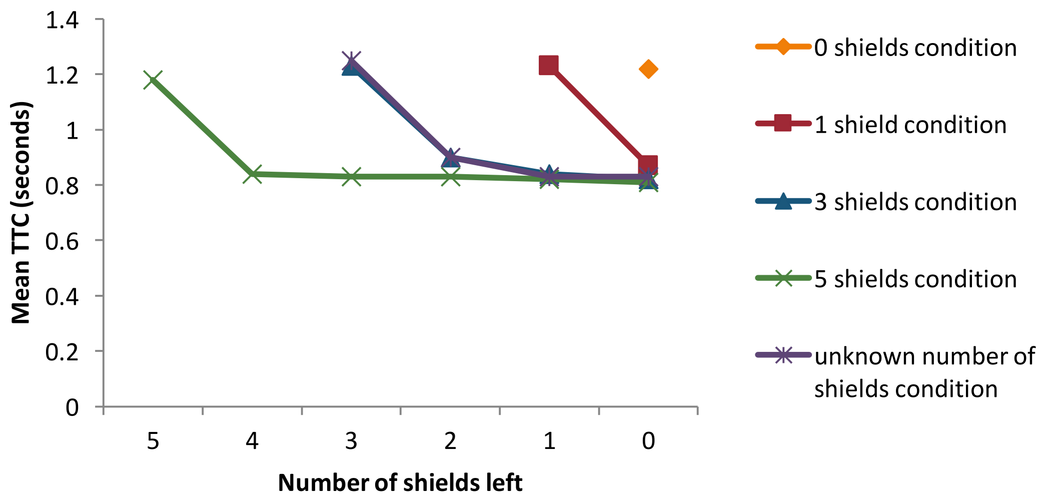

Figure 5. The relationship between the number of shields and the mean TTC.

3.1.4. Mean DCM. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity (χ^2(135) = 873.08, p < .001). Therefore, a Repeated Measures ANOVA with Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the mean DCM (F(6.79, 780.81) = 87.09, p < .001, η_p^2 = .43). However, the post-hoc Bonferroni correction showed that, as before for the mean speed and the mean TTC, only the first shield of every round significantly differed from the rest of the non-starting shields (all ps < .001). The other differences between shields were not significant (all ps > .05). This means that the mean DCM was significantly higher for all first shields compared to any other shield in their condition, but that the non-starting shields did not significantly differ on mean DCM. A graphic representation of the number of shields in relation to mean DCM is shown in Figure 6.

3.1.5. Minimum DCM. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity (χ^2(135) = 168.15, p = .03). Therefore, a Repeated Measures ANOVA with Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the minimum DCM (F(13.24, 1522.57) = 32.59, p < .001, η_p^2 = .22). However, the post-hoc Bonferroni correction showed that, as before for mean speed, mean TTC and mean DCM, only the first shield of every round significantly differed from the rest of the non-starting shields (all p < .001). The other differences between shields were not significant (all ps > .05). This means that the minimum DCM was significantly higher for all first shields compared to any other shield in their condition, but that the non-starting shields did not significantly differ on minimum DCM. A graphic representation of the number of shields in relation to minimum DCM is shown in Figure 7.

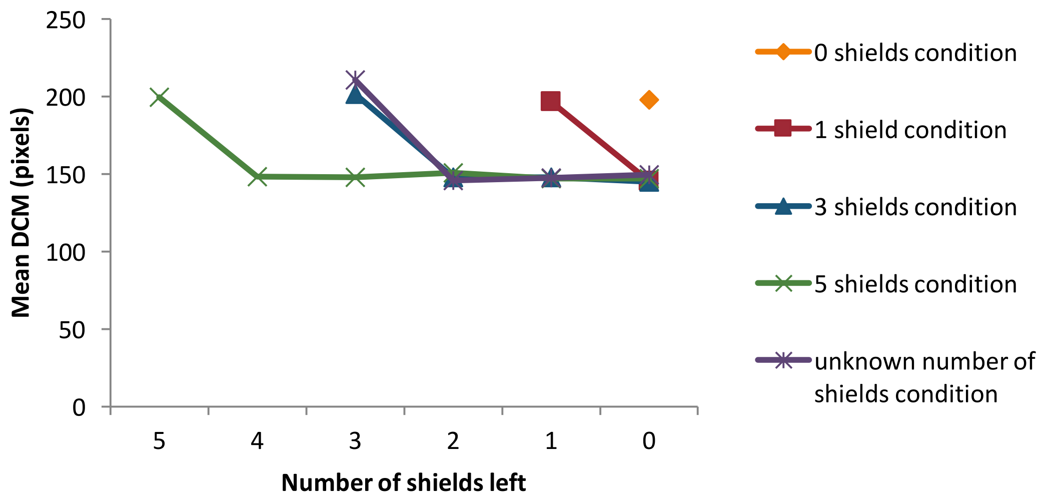

Figure 6. The relationship between the number of shields and the mean DCM.

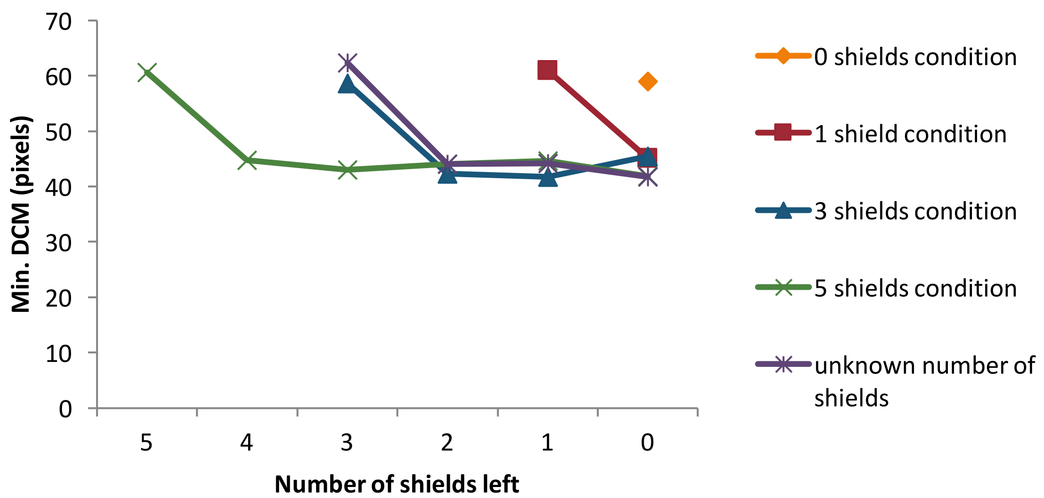

Figure 7. The relationship between the number of shields and the minimum DCM.

3.2. Second hypothesis: Taking a higher risk is related to higher game performance

As for the first hypothesis, the following five measures of risk-taking were used in the current analysis: mean speed, maximum speed, mean TTC, mean DCM and minimum DCM. However, for the current hypothesis the amount of protection, expressed in the number of shields, was considered a sixth measure of risk-taking. Although not freely adjustable by the participants, the number of shields also indicated a certain risk, with less shields left indicating a higher risk. Game performance was measured in distance travelled, and called game score. Unfortunately, the total game score was not normally distributed and thereby violated assumptions of several statistical tests. Therefore, a natural logarithmic transformation was performed before further analysis.

3.2.1. Freely adjustable risk-taking. Pearson's correlation was calculated to assess the relationship between the total game score and the amount of risk that participants were voluntarily taking. Between mean speed and the total game score a weak, negative correlation was found (r = -.21, n = 172, p = .01). Thus, when the mean speed was lower, the total game score was higher. The same was found for the relationship between maximum speed and the total game score (r = -.20, n = 172, p = .01). Between mean TTC and the total game score a weak, positive correlation was found (r = .23, n = 172, p = .01). Meaning that participants with a higher mean TTC had a higher total game score. Between mean DCM and the total game score a moderate, negative correlation was found (r = -.36, n = 172, p < .001). Thus, when the mean DCM was lower, the total game score was higher. However, the correlation between minimum DCM and the total game score was weak and not significant (r = -.13, n = 172, p = .09). Graphic representations of these results are shown in Figures 8 through 12.

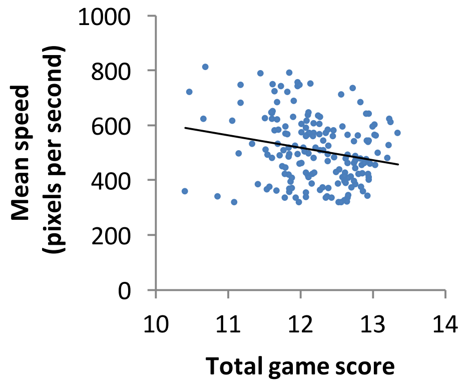

Figure 8. The correlation between the mean speed and the total game score.

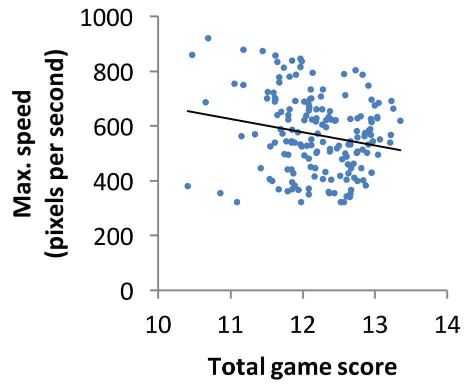

Figure 9. The correlation between the maximum speed and the total game score.

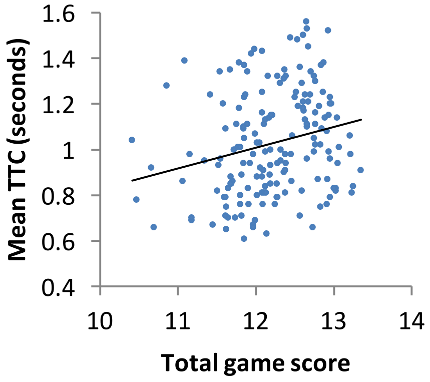

Figure 10. The correlation between the mean TTC and the total game score.

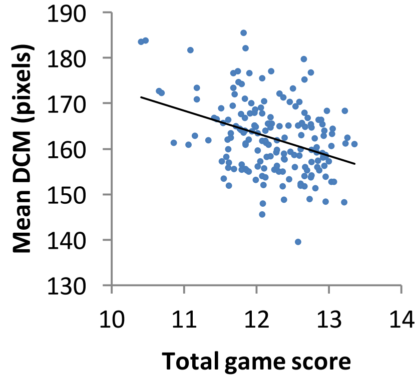

Figure 11. The correlation between the mean DCM and the total game score.

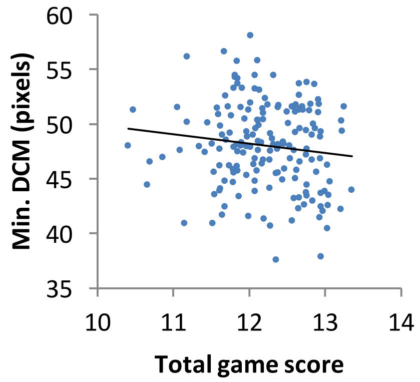

Figure 12. The correlation between the minimum DCM and the total game score.

3.2.2. The number of shields. Mauchly's Test of Sphericity showed a violation of the assumption of sphericity ($\chi^2(135) = 174.02$, $p = .01$). Therefore, a Repeated Measures ANOVA with Greenhouse-Geisser correction was used, which showed that the number of shields had a significant effect on the game score ($F(13.41, 1542.16) = 4.91$, $p < .001$, $\eta_p^2 = .04$). However, the post-hoc Bonferroni correction showed that in the one shield condition there was no significant difference between having or not having a shield ($p = .32$). In the three shields condition the only significant differences emerged between the first shield and when there were no shields left ($p = .02$). The other differences in this round were not significant (all *ps* > .05). In the five shields condition the first and the second shield significantly differed from one another ($p = .01$). No other differences were found in this condition (all *ps* > .05). In the unknown number of shields condition no significant differences were found (all *ps* > .05). This means that the game score was only significantly higher for the first shield of the three shields condition compared to having no shields left, and the first shield of the five shields condition compared to the second shield. A graphic representation of the number of shields in relation to the game score is shown in Figure 13.

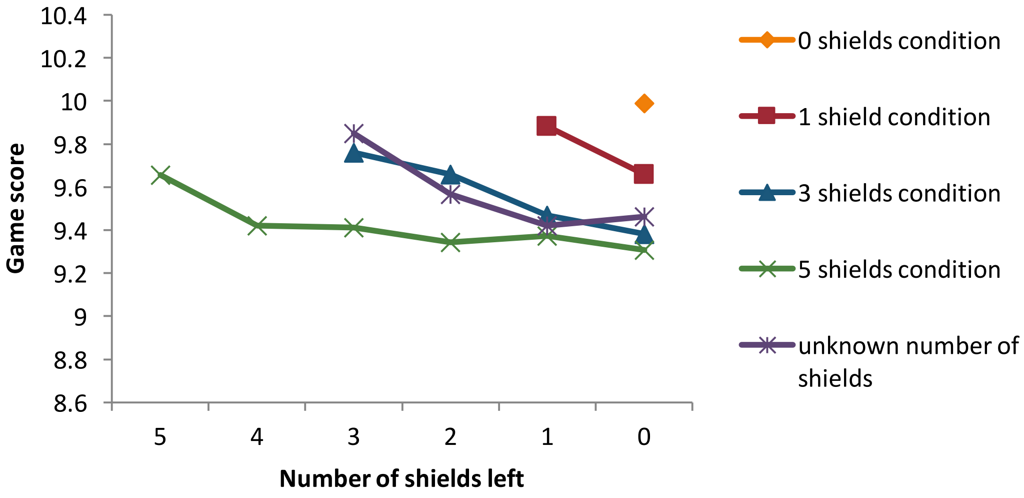

Figure 13. The relationship between the number of shields and the game score.

3.3. Third hypothesis: Participants in a high positive affective state will take more risk than participants in a low positive affective state

A median split of the positive affect score (*Mdn* = 26.00) was performed to divide the participants into a *low positive affect group* (*N* = 78) and a *high positive affect group* (*N* = 94). Independent samples t-tests were executed to compare these groups, and Pearson's correlation was calculated to assess the overall relationship between positive affect and risk-taking.

3.3.1. Mean speed. The mean speed was higher in the low positive affect group (*M* = 527.78, *SD* = 136.80) than in the high positive affect group (*M* = 494.60, *SD* = 106.06), meaning that the low positive affect group took more risk. However, this difference was not significant ($t(170) = 1.79$, $p = .08$). The overall correlation between positive affect and mean speed was very weak and not significant ($r = -.05$, $n = 172$, $p = .56$). Graphic representations of these results are shown in Figure 14 and Figure 15 respectively.

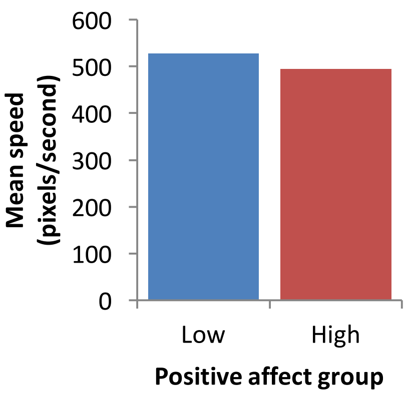

Figure 14. The difference in mean speed between the positive affect groups.

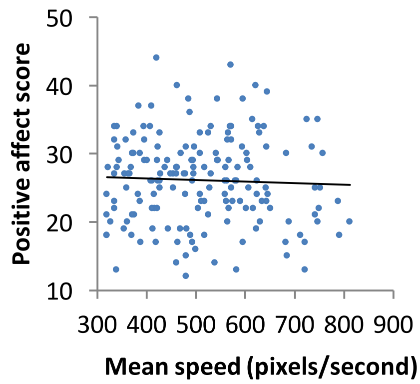

Figure 15. The correlation between the positive affect score and mean speed.

3.3.2. Maximum speed. The maximum speed was higher in the low positive affect group (*M* = 583.29, *SD* = 161.82) than in the high positive affect group (*M* = 555.59 *SD* = 123.26), meaning that the low positive affect group took more risk. Again, this difference was not significant ($t(170) = 1.27$, $p = .21$). The overall correlation between positive affect and maximum speed was very weak and not significant ($r = -.02$, $n = 172$, $p = .83$). Graphic representations of these results are shown in Figure 16 and Figure 17 respectively.

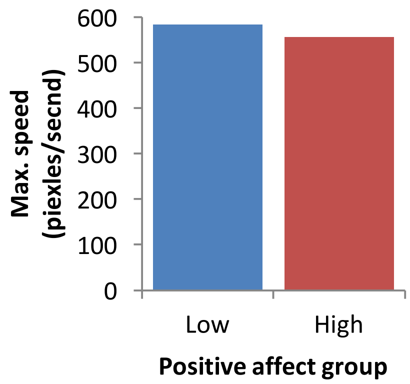

Figure 16. The difference in maximum speed between the positive affect groups.

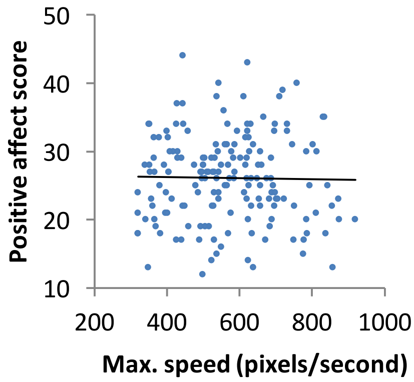

Figure 17. The correlation between the positive affect score and maximum speed.

3.3.3. Mean TTC. The mean TTC was lower in the low positive affect group (*M* = 1.00, *SD* = 0.24) than in the high positive affect group (*M* = 1.05, *SD* = 0.22), meaning that the low positive affect group took more risk. However, this difference was not significant ($t(170) = -1.34$, $p = .18$). The overall correlation between positive affect and mean TTC was very weak and not significant ($r = .01$, $n = 172$, $p = .87$). Graphic representations of these results are shown in Figure 18 and Figure 19 respectively.

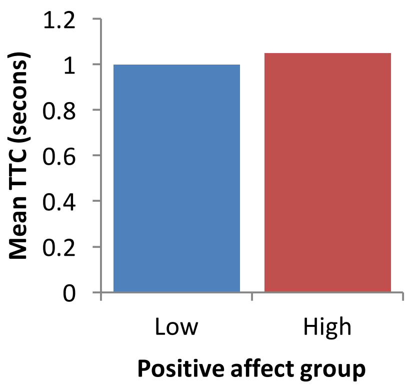

Figure 18. The difference in mean TTC between the positive affect groups.

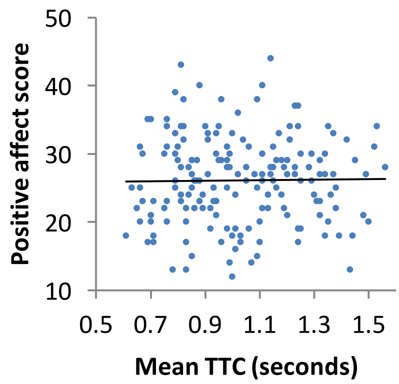

Figure 19. The correlation between the positive affect score and mean TTC.

3.3.4. Mean DCM. The mean DCM was lower in the high positive affect group (M = 162.10, SD = 7.55) than in the low positive affect group (M = 162.69, SD = 8.56), meaning that the high positive affect group took more risk. However, this difference was not significant ($t(170) = 0.48$, $p = .63$). The overall correlation between positive affect and mean DCM was weak and not significant ($r = -.12$, $n = 172$, $p = .11$). Graphic representations of these results are shown in Figure 20 and Figure 21 respectively.

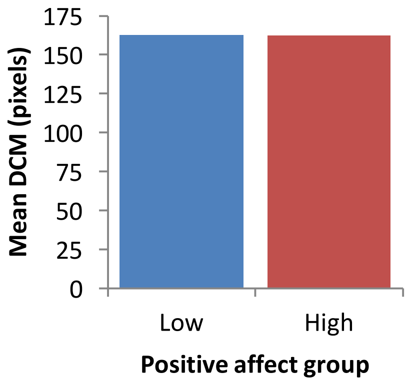

Figure 20. The differences in mean DCM between the positive affect groups.

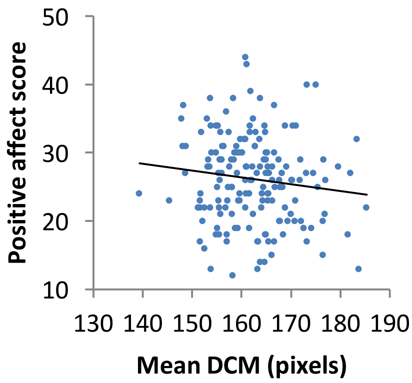

Figure 21. The correlation between the positive affect score and mean DCM.

3.3.5. Minimum DCM. The minimum DCM was lower in the high positive affect group (M = 47.63, SD = 3.95) than in the low positive affect group (M = 48.50, SD = 3.84), meaning that the high positive affect group took more risk. Again, this difference was not significant ($t(170) = 1.46$, $p = .15$). The overall correlation between positive affect and minimum DCM was very weak and not significant ($r = -.08$, $n = 172$, $p = .29$). Graphic representations of these results are shown in Figure 22 and Figure 23 respectively.

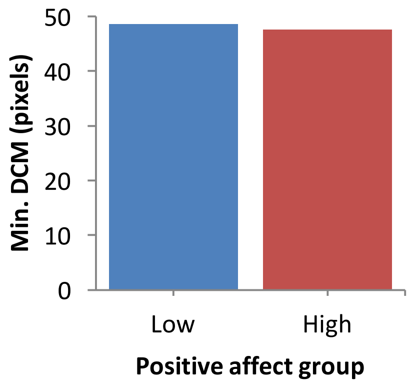

Figure 22. The differences in minimum DCM between the positive affect groups.

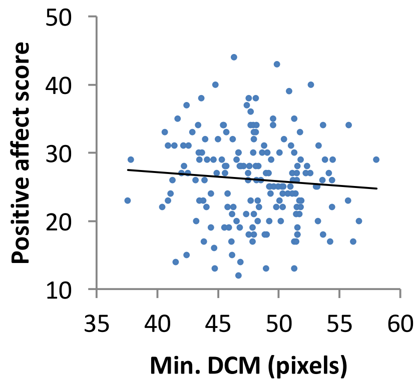

Figure 23. The correlation between the positive affect score and minimum DCM.

3.4. Fourth hypothesis: Participants in a high positive affective state will perform better than participants in a low positive affective state

As in the analysis of the second hypothesis, the natural logarithmic transformation of the total game score was used in the analysis of the current hypothesis. Also, as in the analysis of the third hypothesis, the participants were divided into a low positive affect group ($N = 78$) and a high positive affect group ($N = 94$) by using the median split of the positive affect score ($Mdn = 26.00$). An independent samples t-test was performed to compare the two groups. In addition, Pearson's correlation was calculated to assess the relationship between the positive affect score and the total game score.

The high positive affect group had a higher total game score ($M = 12.31$, $SD = .55$) than the low positive affect group ($M = 12.10$, $SD = .59$), meaning that the high positive affect group performed better on the video game. This difference was significant ($t(170) = -2.45$, $p = .02$). The overall correlation between the positive affect score and the total game score was weak, but significant ($r = .18$, $n = 172$, $p = .02$). Meaning that when the positive affect score was higher, the total game score was also higher. Graphic representations of these results are shown in Figure 24 and Figure 25 respectively.

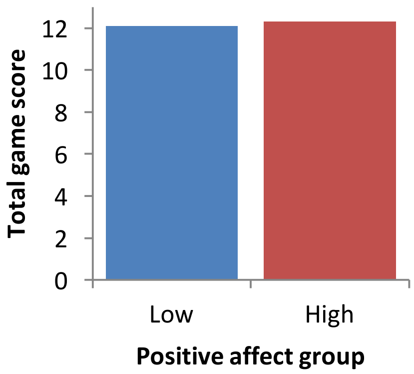

Figure 24. The differences in total game score between the positive affect groups.

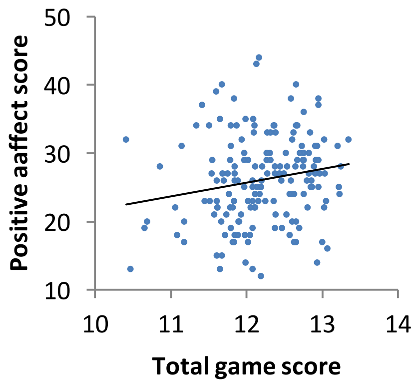

Figure 25. The correlation between the positive affect score and the total game score.

4. Discussion

The main objective of the current study was testing Wilde's (1982) Risk Homeostasis Theory. To do so, a video game was created in which participants flew a spaceship through a meteor shower. The spaceship was equipped with shields that would protect it, but would be lost during a collision. Additionally, the effect of affect on RHT was investigated. A total of four hypotheses were tested in the current experiment. In this chapter, these hypotheses will be separately discussed, followed by suggestions to improve the current study and an overall conclusion.

4.1. First hypothesis: Participants will behave as predicted by RHT

It was hypothesized that participants would behave as predicted by Wilde's theory, meaning that participants would adjust their behavior to compensate for added or reduced safety, thereby balancing their perceived and target level of risk. In the current experiment, this would have entailed that participants decreased the amount of risk they were taking whenever their spaceship lost a shield.

Surprisingly, the results show that participants significantly increased the amount of risk they were taking after they had lost the first shield. Apart from one exception this is true for all conditions on all five measures of risk-taking: mean speed, maximum speed, mean TTC, mean DCM and minimum DCM. Only in the one shield condition there was no significant difference on maximum speed between having or not having a shield. At first glance, these results seem to point towards the opposite of what was expected, thus a reversed RHT effect. However, a truly reversed RHT effect would entail that participants kept increasing the amount of risk they were taking. This was not the case. The differences between the remaining shields are rather small and, apart from one minor exception, not significant. This is true for all measures of risk-taking.

The fact that the first shield of every round systematically differs from the rest of the shields is likely caused by a design flaw. Each round of the game started with the minimum speed of 320 pixels per second which, in steps of 50 pixels per second, could be increased to 920 pixels per second. It is reasonable to assume that the starting speed was perceived as too slow, and that participants needed time to accelerate the spaceship to a comfortable speed. This also explains the same trend in TTC and DCM, as speed is inherently related to the time participants had to dodge meteors. In addition, it must be noted that the meteors entered the screen from the right side, thus the farthest away from the spaceship. Meaning that TTC and DCM were automatically higher, thus less risky, during the first few moments of all rounds.

Due to these design flaws the data from all first shields and the zero shields condition are systematically biased, and will be disregarded during the further discussion of the results.

The remaining data show that participants did not significantly alter their behavior once they had found a supposedly comfortable flying speed, TTC and DCM to dodge meteors. This is generally true for all measures of risk-taking, apart from the mean speed in the unknown number of shields condition. In this condition, the mean speed for the second shield was also significantly lower than the speed during all following shields. This difference again hints towards a reversed RHT effect. However, when taking Figure 3 through Figure 7 into account, it can be observed that all conditions gradually advance towards roughly the same value, and that this is true for all measures of risk-taking. This means that participants progressed towards the same level of risk-taking, regardless of the number of shields a round started with.

It is evident that the data from this experiment do not support the first hypothesis. However, this does not automatically mean that they refute RHT. They merely point out the flaws of the current experiment and the difficulties of testing RHT in an experimental setting. One possible explanation for why the number of shields, thus the amount of protection, did not influence behavior could be that the current experiment failed to realistically portray risks and benefits. Participants knew that they would receive an extra reward if they had the best score, which would be a real-life benefit. However, they also knew that they would receive a set amount of money for their participation, regardless of their performance, and that losing a shield had no real-life consequences. Therefore, it is presumed that the incentive to perform and the risk of losing something were not high and realistically enough for risk homeostasis to occur. Of course, another explanation for the results could be that RHT is not true. The problem with this conclusion is that the data from the current experiment do not refute RHT, since it is impossible to prove that the results are not due to, for example, a change in the target level of risk or the deficient portrayal of risk and benefits. Like Adams (1988) and Hoyes and Glendon (1993) already pointed out, it seems impossible to refute RHT, due to its metaphysical and vague formulation.

4.2. Second hypothesis: Taking a higher risk is related to higher game performance

Furthermore, it was assumed that people take more risk in order to maximize their benefits. Therefore, it was investigated if taking more risk is truly beneficial for reaching one’s goal, thus if taking a higher risk is related to higher game performance. Translated to the current experiment this means that taking more risk in terms of mean speed, maximum speed, mean TTC, mean DCM and minimum DCM is related to a higher total game score. Additionally, the influence of the number of shields left was investigated. Although not freely adjustable, a lower number of shields indicated a higher risk.

When analyzing the data from the number of shields, only two significant differences were found. One in the three shields condition between the first shield and when there were no shields left, and the other in the five shields condition between the first and the second shield. However, due to the bias regarding the data collected at the beginning of each round, which is discussed in paragraph 4.1., these results are unreliable and should be neglected. The remaining data show that the number of shields had no significant impact on the game score. This finding is not surprising, since the data from the first hypothesis also showed that the number of shields had no significant effect, which is likely due to the inadequate portrayal of risks and benefits.

The results regarding the five measures of risk-taking are more interesting. It was found that taking a lower risk is weakly but clearly significantly related to a higher game performance on the measures mean speed, maximum speed and mean TTC. On the other hand, the data for mean DCM show that higher risk-taking is moderately related to a higher game performance. This result is also significant. The data for minimum DCM points in the same direction as the data for mean DCM, but the correlation is not significant. One might be inclined to conclude that the benefits of high risk-taking depend on the kind of risk. However, these contradictory results are more likely due to the relationship between the measures of risk-taking in this experiment. When participants navigated their ship more slowly through the meteor shower, they had more time to oversee the situation and to estimate the distance needed to safely pass meteors. It is plausible that the slower participants could more precisely plan and execute their maneuvers, making it unnecessary for them to excessively deviate from their set route. Thus, it seems more likely that the higher risk on the mean DCM measure is a byproduct of a lower risk on speed, than a potential cause of a higher game score. This reasoning makes speed the most important measure of risk-taking, which is plausible since speed is inherently related to both TTC and DCM.

Clearly, the data do not support the hypothesis that taking a higher risk is related to higher game performance. In fact, the data show the opposite: taking a lower risk is related to higher game performance. It seems that, in the current experiment, the best strategy for maximizing one’s benefits was voluntarily taking a low risk. However, it must be noted that this assumption is based on correlational research, and that correlation does not necessarily imply causation. Also, the correlations found here represent rather small effect sizes. Which means that other factors might have a stronger relationship with the measured parameters. Since the current experiment revolved around a video game, gaming skill or reaction time are likely contenders. Furthermore, the assumption that the best strategy for maximizing one's benefits was taking a low risk can only be true if the participants' goal was reaching the highest possible game performance. The problem here is, that their actual goal is unknown to the researchers. It might be the highest possible game performance because participants wanted to win the extra reward, but it could also be finishing the experiment as quickly as possible, regardless of the game performance, to receive the participation reward. There are many more possible goals, and all have their own subjective measure of risk. Therefore, one cannot be sure that the current experiment measured the right parameters.

4.3. Third hypothesis: Participants in a high positive affective state will take more risk than participants in a low positive affective state

The third hypothesis investigated in this study was that participants in a high positive affective state will take more risk than participants in a low positive affective state. It was argued that affect plays a role in establishing the target level of risk in RHT, by modulating one's perception of costs and benefits. Therefore, participants in different affective states were expected to take different amounts of risk.

The results show inconsistencies that are similar to the ones observed in the data from second hypothesis: on the measures mean speed, maximum speed and mean TTC, participants in the low positive affect group took more risk than participants in the high positive affect group, while the data for mean DCM and minimum DCM show the opposite. Again, these inconsistencies are likely due to the relationship between speed and DCM, which is discussed in paragraph 4.2. However, the differences found in the current analysis are extremely small and not significant. Meaning that the two affect groups did not significantly differ on risk-taking. The general directions of the independent samples t-test are supported by the overall correlations between positive affect and the five measures of risk-taking. Nevertheless, these correlations are also very small and not significant. This means that there is no significant relationship between positive affect and risk-taking in the current experiment.

The fact that none of the results are significant might be due to the participants' target levels of risk being too similar. It was argued that affect influences the perception of risk and benefits and therefore the target level of risk. However, due to risk and benefits not being portrayed realistically enough in this experiment, the range of target levels might have been very limited, causing no significant differences in behavior. Unfortunately, it is impossible to measure one's target level of risk to test this assumption. Also, since the participants' actual goals and subjective measures of risk are unknown, the current experiment might have failed to measure the right parameters. Another explanation for the results could be that affect does not influence the target level (in such a limited amount of time), or that there is no target level because RHT is not true.

4.4. Fourth hypothesis: Participants in a high positive affective state will perform better than participants in a low positive affective state

The last hypothesis investigated in this study was that participants in a high positive affective state will perform better than participants in a low positive affective state. This hypothesis is based on the Broaden-and-Build Theory of Positive Emotions by Fredrickson and Branigan (2005), which states that emotions influence the attentional scope. In turn, a lack of attention affects the perceived level of risk (Wilde, 1982). Therefore, it was argued that affect would influence the perceived level of risk. Since the perceived level of risk cannot be measured, the current experiment measured performance outcomes.

The results support the hypothesis by showing that the high positive affect group had a significantly higher total game score than the low positive affect group. The correlation between positive affect and the total game score supports this by showing that, overall, a higher positive affect score is related to a higher total game score. However, it should be noted that this correlation represents a small effect size and that, as discussed in paragraph 4.2., other factors likely have a stronger relationship with performance. Overall, the current results are in line with the Broaden-and-Build Theory of Positive Emotions, nevertheless, it cannot be concluded that affect influences the perceived level of risk. Especially since the current study did not find support for the homeostatic mechanism and suffers from a variety of methodological shortcomings that might have distorted the results. It is very well possible that the observed differences are only due to the broadened attentional scope. Because the perceived level of risk itself cannot be measured, the current experiment can only hint towards the influence of affect.

4.5. Potential improvements to the current experiment

The following suggestions should be considered if the current study is replicated. More realistic portrayals of risks and benefits should be utilized. This could be achieved by integrating the monetary reward for participation in the gameplay. For example, by telling participants that they will lose a certain amount of money whenever they lose a shield, and that they can gain money by reaching the destination quicker. It could also be considered to not automatically provide the participants with shields, but to let them decide if they want to purchase them. This would add a new dimension of risk-taking, however, it would also limit the control over the experimental conditions.

Furthermore, the design flaw regarding the first shield of every round could be repaired by starting each round with a few seconds of immunity, in which the spaceship cannot be destroyed, before logging the game data. Then, participants would have time to experiment with different speeds and to accelerate to a comfortable one. Additionally, the meteors would be scattered over the entire screen by the time the game starts logging data, also preventing the bias in TTC and DCM.

Another problem of the current experiment was that it was unclear what the participants' goals were. Therefore, the corresponding risks were also not clear. The more realistic portrayal of risks and benefits would probably reduce this problem. However, verifying that the right parameters are measured could be done by asking the participants about their goals when the experiment has ended.

4.6. Conclusion

The current study was an attempt to research RHT and the effect of affect on two of its major components; the target and perceived level of risk. No new evidence for or against RHT was found. However, the current experiment highlights the problems of testing RHT in general. The largest problem is that some of the theory's major components are unmeasurable, and that one is forced to rely on behavioral outcomes like risk-taking and performance. This is not only true for the current experiment, but also for the current state of research technology. Without measuring RHT's essential components there is no way to verify or refute the theory completely. It is undeniable that the current experiment suffers from several methodological shortcomings, however, even after straightening those out the experiment could only grasp at what is happening in the black box that is RHT. Further research on RHT is encouraged, however, this will only be fruitful if RHT is formulated more clearly and it becomes possible to measure its major components.

5. References

Adams, J. G. U. (1988). Risk homeostasis and the purpose of safety regulation. *Ergonomics, 31*(4), 407-428. doi:10.1080/00140138808966688

Baniela, S. I., & Ríos, J. V. (2010). The Risk Homeostasis Theory. *The Journal of Navigation, 63*(4), 607-626. doi:10.1017/S0373463310000196

Evans, L. (1986). Risk Homeostasis Theory and Traffic Accident Data. *Risk Analysis, 6*(1), 81-94. doi:10.1111/j.1539-6924.1986.tb00196.x

Finucane, M. L., Alhakami, A., Slovic, P. & Johnson, S. M. (2000). The Affect Heuristic in Judgments of Risks and Benefits. *Journal of Behavioral Decision Making*, *13*(1), 1-17. doi:10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion, 19*(3), 313-332. doi:10.1080 /02699930441000238

Gasper, K. (2004). Do you see what I see? Affect and visual information processing. *Cognition and Emotion, 18*(3), 405-421. doi:10.1080/02699930341000068

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and Global Versus Local Processing of Visual Information. *Psychological Science, 13*(1), 34-40. doi:10.1111/1467-9280.00406

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*(1-3), 49-62. doi:10.1016/0925-7535(96)00005-7

Hoyes, T. W., & Glendon, A.I. (1993). Risk homeostasis: issues for future research. *Safety Science, 16*(1), 19-33. doi:10.1016/0925-7535(93)90004-W

Isen, A. M., & Patrick, R. (1983). The Effect of Positive Feelings on Risk-taking: When the Chips Are Down. *Organizational Behavior and Human Performance, 31*(2), 194-202. doi:10.1016/0030-5073(83)90120-4

Jackson, J. S. H., & Blackman, R. (1994). A Driving-Simulator Test of Wilde's Risk Homeostasis Theory. *Journal of Applied Psychology, 79*(6), 950-958. doi:10.1037 /0021-9010.79.6.950

O'Neill, B., & Williams, A. (1998). Risk homeostasis hypothesis: a rebuttal. *Injury Prevention, 4*(2), 92-93. doi:10.1136/ip.4.2.92

Rowe, G., Hirsh, J. B., & Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences, 104*(1), 383 -388. doi:10.1073/pnas.0605198104

Siegel, E. H., & Stefanucci, J. K. (2011). A little bit louder now: Negative affect increases perceived loudness. Emotion, 11, 1006-1011. doi:10.1037/a0024590

Slovic, P. (2010). The Psychology of Risk. *Saúde e Sociedade, 19*(4), 731-747. doi:10.1590 /S0104-12902010000400002

Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, Risk, and Decision Making. *Health Psychology, 24*(4), S35-S40. doi:10.1037/0278-6133 .24.4.S35

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. doi:10.1037/0022-3514.54.6.1063

Wilde, G. J. S. (1982). The Theory of Risk Homeostasis: Implications for Safety and Health. *Risk Analysis, 2*(4), 209-225. doi:10.1111/j.1539-6924.1982.tb01384.x

Wilde, G. J. S. (1998). Risk homeostasis theory: an overview. *Injury Prevention, 4*(2), 89-91. doi:10.1136/ip.4.2.89

Wilde, G. J. S. (2014). Risk homeostasis in the laboratory. In *Target Risk 3: Risk Homeostasis in Everyday Life* (pp. 119-133). Retrieved from http://www .riskhomeostasis.org

Yuen, K. S. L., & Lee, T. M. C. (2003). Could mood state affect risk-taking decisions? *Journal of Affective Disorders, 75*(1), 11-18. doi:10.1016/S0165-0327(02)00022-8