An investigation of risk compensation and the influence of gain-framed and loss-framed messages on risk behavior in a laboratory-based video game experiment

Master Thesis in Occupational Health Psychology

by Marius Deckers

Studentnumber: s1576992

marius-deckers@web.de

June 29th, 2015

Supervisor: Jop Groeneweg, PhD

Faculteit der Sociale Wetenschappen

Leiden University

Content:

Abstract

This thesis discusses the concepts of risk behavior and Risk Homeostasis Theory as proposed by Wilde (1982a). While additional theory is presented about gain- and loss-framed messages and their possible effects on behavior, risk behavior is discussed both theoretically and practically in the context of negative and positive consequences of behavior and in light of recent research. Both notions were combined into an experiment that allowed for objective measurements of risk behavior. This experiment tested the influences of *a)* varying degrees of objective safety and *b)* of gain- and loss-framed messages highlighting the consequences of actions on risk behavior. While gain- or loss-framed messages both showed no statistically significant influence on risk behavior in comparison to a control group, a higher amount of objective safety was compensated for by participants through more risky behavior, albeit not in all variables used to measure risk behavior. Potential flaws in the experimental manipulations and the implications of these finding for the future of RHT research and accident prevention are discussed.

1 Introduction

1.1 Overview of risk, risk behavior and homeostasis

It is common knowledge that part of the human behavior repertoire is taking risk. From an evolutionary perspective, it made sense for our ancestors to engage in potentially dangerous activities when the reward was high. As a hunter, hunting a large animal will provide an entire family or clan with food and other important supplies for a long period of time; however, large animals can also pose a considerable danger to the group of hunters. Therefore, taking the risk of being injured during a hunt was sometimes a necessity if food was scarce. In the present day, regardless of interindividual differences in the amount of risk people are generally willing to take, everybody exhibits behavior at least at a few points in their life that can be considered "risky". But what exactly is risk and why do people take it?

There are multiple definitions of risk in circulation and the following one has been created for the purpose of this thesis:

"Risk is a subjective measurement for the probability of a negative consequence occurring because of any action or decision. The measurement is subjective because every individual perceives the weight of negative consequences and its relation to positive consequences differently. Risk is therefore related to the uncertainty of outcome as without uncertainty, a decision never bears risk."

This definition is broad in nature to enable it to be generalized across multiple domains. The connection to the uncertainty of outcome is made out of logical necessity – if the outcome of an action or decision is absolutely certain and therefore without alternatives, it cannot be considered risky. It should be noted that absolute certainty of outcome is very likely unattainable in real situations and is more of theoretical nature, as even for simple actions, the long-term consequences are not calculable in a reliable manner. Regardless, the outcome of an action is a large factor in the decision process for carrying out that action.

It is safe to assume that a negative consequence of a behavior is something people want to avoid. However, why is it the case that humans exhibit behavior that can potentially have negative consequences? Why do they take risks? The obvious answer would be that they expect a positive outcome of any kind that outweighs the negatives – however, the nature of these potential gains is not always inherently obvious. While it seems intuitive to say that humans are able to judge the weight of positive vs. negative outcomes of a behavior, the fact that some people are willing to take a lot more risk than others makes this matter slightly more complicated; are some people just better at "objectively" judging risk and potential loss and gains? A better explanation lies within the context of sensation-seeking. It has been shown that individuals who more strongly seek sensation also show more risky behavior (Horvath & Zuckerman, 1992). Further, high sensation-seekers tend to estimate risk as lower than low sensation-seekers, even if both groups have never participated in the (risky) activity in question. Adding to these findings, research has also demonstrated that the relationship between sensation-seeking and risky behavior is mediated by the judgment of cost and benefit (Maslowsky *et al.*, 2011). Specifically, individuals higher in sensation-seeking assign more value to the potential benefits of a behavior and less (negative) value to the potential costs in comparison to individuals low in sensation-seeking, for whom the inverse relationship for judging gains and losses has been found. Additionally, valuing the potential benefits higher than the costs is a considerable predictor for engaging in a risky behavior. Therefore, sensation-seeking seems to play an important role in the explanation as to why humans take risks. However, when assessing the questions how and when humans take risk, one needs to look into the mechanisms behind risk assessment and risk taking behavior.

Originally, the term *homeostasis* was not created to explain human risk behavior. Homeostasis originates in the Greek language and is a combination of *homœos*, meaning "similar", and *stasis*, meaning "standing still". The term describes the process within a system that allows its conditions to remain constant over time through changes in variables that, in the grand scheme, cancel each other out so that the overarching state of the system remains the same. Examples for such homeostatic systems are body temperature (only in endothermic animals such as mammals and birds), which remains constant regardless of varying outside temperatures and other conditions, as well as regulation of the blood pH at 7.365. Even though the term was coined in 1926, if was not until more than 50 years later that it was used in combination with risk behavior.

In 1982, Wilde developed the theory of *risk homeostasis* (Wilde, 1982a). Originally, risk homeostasis theory (from here on abbreviated as RHT) was supposed to serve as an overarching explanatory framework for traffic accidents and their causation. Wilde identifies risk taking as one of the primary causes for accidents and proposes that a person engaging in driving a vehicle (or in a comparable task) at every moment evaluates two distinct risk levels: the state of the perceived risk level and the state of desired (or target) risk level. According to Wilde, in case of a discrepancy between these two risk levels, a person adjusts their behavior to match them again. This indicates that, in case the perceived risk level is lower than the desired risk level, the person will behave more recklessly, while switched risk levels cause the person to adjust their behavior into the other direction and be more cautious. The main point of Wilde's proposition is that the desired level of risk in a population remains essentially constant over time and that therefore safety regulations or tools such as seatbelts have, at most, a temporary effect on the overall number and severity of traffic accidents. In the example of seatbelts, while they do increase the safety of the vehicle being driven, they do not increase the driver's desire to be safe, resulting in a behavioral change caused by the decrease in perceived risk level such as driving faster or more recklessly, ultimately canceling out the added safety effect of the seatbelt. This is generalized to all safety standards, meaning that although vehicles become safer, accident rates show no measureable deviation from consistency over time, which Wilde cites evidence for. Originally, the theory was labeled risk *compensation*, but to align the name with the findings just described, the term homeostasis was deemed to fit better. In RHT, accident rates are compared to a thermostat, where the only way to change the resulting room temperature is to change the desired temperature in the settings. All other changes merely cause the thermostat to adapt, trying to keep the room temperature at the level it was assigned to keep it at. While this comparison does have the shortcoming of comparing a simple technical device to the dynamic and reactive variables of traffic, it has an important implication for the overall reduction of occurrence and severity of traffic accidents: Wilde ultimately proposes that the only way to reliably reduce traffic accidents and the associated death rates is to reduce the amount of risk drivers are willing to take, or, in other words, increase their desire for safety (in the thermostat analogy, this would equal to lowering the desired temperature on the thermostat). The underlying processes are outlined in the model below.

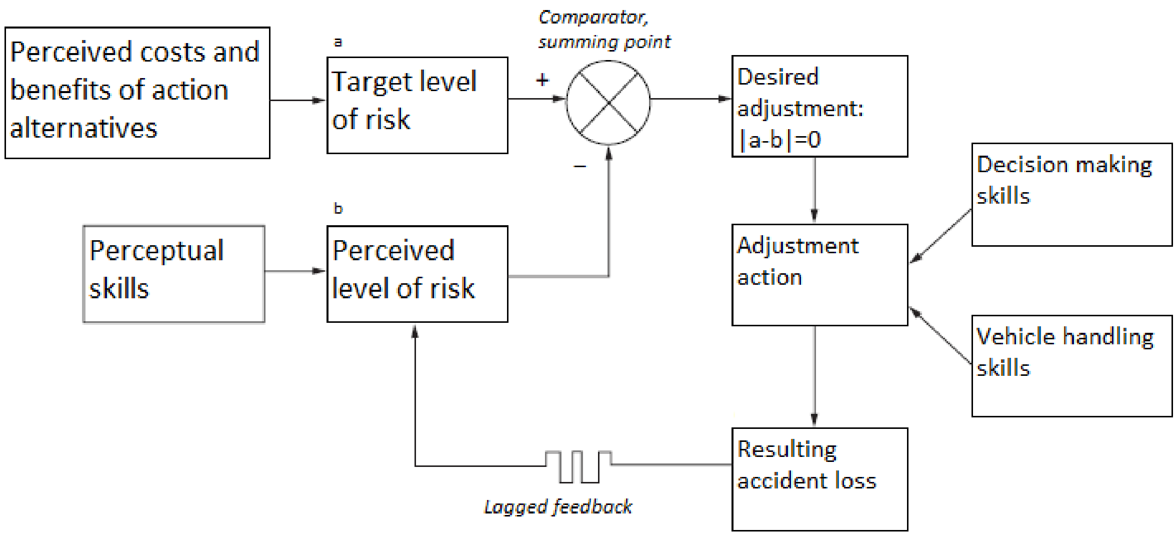


Figure 1: Wilde's original model created to explain the homeostatic mechanism is displayed as a slightly simplified version.

The question arises whether RHT can be applied to domains other than traffic. It has long been hypothesized that humans are characterized by an optimal level of psychophysiological arousal (for the initial publications on this idea, see Leuba, 1955 and Hebb, 1955). It can therefore be hypothesized that taking considerable risks is a means to achieve this optimal arousal level, which would in turn indicate that RHT is applicable to areas other than traffic and is, in fact, quite a universal theory.

Historically, RHT has not been developed much further theoretically; however, it has been both tested in simulated environments and evaluated in real situations. An early example is the study of Jackson & Blackman (1994), which aimed to investigate whether the propositions made by RHT could be confirmed in a driving simulator test. RHT predicts that only changes in the amount of risk people are willing to take will have a (lasting) effect on accident rates. In this regard, Jackson & Blackman chose either high or low accident costs as a manipulation, expecting accident costs to influence the amount of risk people are willing to take through changes in the costs- and benefits-analysis. Additionally, a high or low speed limit and a high or low fine for speeding were used as manipulations with the expectation that both should not have an influence on overall accident occurrence as they do not influence the desired (target) level of risk. The results of this experiment confirmed the hypotheses made and, upon further data investigation, revealed that increasing the accident costs did indeed decrease accident frequency; however, it had no influence on the speed chosen by the drivers. This might indicate that increasing accident costs might heighten alertness and salience of accidents, effectively increasing focus and attention while driving at higher speed, making accidents caused by gaps in attention less likely to occur.

Another study in this area was conducted by Glendon *et al.* (1996), which investigated RHT in the context of utility, with the question if taking risk and the homeostatic process exist either "for their own sake" or only if there is some form of utility for the risk-taker. It was found that risk homeostasis might have to be viewed in the context of sensation-seeking, which of course can be seen as a form of utility as well. Additionally, Glendon and colleagues observed that risk *compensation* can occur in the very short term (RHT generally predicts homeostatic processes to take place over *years*, through lagged feedback) and that those short-term adjustments to new risk levels might be over-compensatory. The authors implied that future research needs to identify the exact pathways through which risk homeostasis takes place and "block" them experimentally, answering the question whether risk homeostasis disappears or finds another utility or new pathway through which it will manifest itself. This approach has not been taken experimentally to the present day.

Although more theoretical in nature, RHT has been proposed to have major implications for and contributions to real-life situations. As the theory predicts that perceived level and target level of risk are a major influence on risk behavior, it has been proposed (Wilde, 1989) that effective and lasting accident reduction can only be achieved through a reduction of the target level of risk in the population as a whole. This corresponds to lowering the target level of risk in every individual, which, viewed from the opposite perspective, means increasing everybody's desire for safety, which is assumed to lead to the same result. Accordingly, another repeatedly occurring claim (Simonet & Wilde, 1997) was that making activities objectively safer (for example through the addition of seatbelts or other safety equipment) has no lasting effects on actual accident occurrence as it does not influence people's desire to be safe. According to Wilde (2005), providing a safety- incentives program would be fruitful: for drivers to manage their risk better, their recognition of risk must improve and their willingness to accept risk must decline. While the former factor is difficult to influence, the latter has been successfully reduced in programs involving incentives (for example monetary) to keep the amount of risky actions while driving to a minimum. Wilde additionally proposes that incentives are inherently better than rewards as incentives can influence behavior *before* they are given.

Additional support for RHT "from the field" was provided as well (Aschenbrenner & Biehl (1994), as cited by Simonet & Wilde, 1997). In Munich, taxicabs started to get equipped with ABS (anti-lock brakes), which is a system that keeps the ability to steer the car intact even in the event of a full application of the brake. In line with RHT's predictions, the taxi drivers who had their vehicles equipped with such a system changed their driving behavior to compensate for the added safety provided, resulting in the accident rates of taxicabs staying constant over time. These findings were confirmed in another country as well, with the added realization that drivers used the new ABS to increase speed instead of decreasing braking distance (Grant & Smiley (1993), as cited by Simonet & Wilde, 1997). Finally, Simonet and Wilde (1997) introduced the business cycle as another influence on accident rates, providing evidence that traffic death rates increase in times of economic prosperity while they decrease in times of financial crisis or recession. This is viewed as a confirmation for the underlying processes of RHT: in times of critical financial situations, the costs of accidents (e. g. car repairs) receive considerably more weight, while the perceived potential (monetary) gains made through speeding and other risky behavior are reduced. Lastly, although RHT was created for car traffic, a confirmation for its mechanics was found elsewhere (Baniela & Ríos, 2010). Despite continuous improvements of safety standards and regulations in ship navigation, the number of observed shipping casualties has not dropped measurably. The explanation given by the authors lies partially in the payment system commonly used in ship transportation: It is common practice to pay Captains based on how quickly they can reach their destination. In this light, it is understandable that new safety standards cause a change in perceived level of risk, which, combined with the monetary incentive to rush, causes Captains to go faster than usual because they are under the impression that their ships are now safer than before, ultimately going fast and careless enough to cancel out the added safety effect of new measures or regulations. This kind of compensatory behavior is in line with the proposed mechanics of RHT.

As with many scientific theories, RHT is not without its critics. One very straightforward but thorough analysis of existing traffic data (Evans, 1986) came to the conclusion that all these data do not support RHT and appear to be refuting it in part or as a whole. The proposed equilibrium of fatalities has never been found in many differently coded traffic accident data and even an analysis of *all* recorded fatalities (not only in traffic) since 1900 did not reveal a homeostatic trend. Additionally, contrary to the predictions of RHT, different types of roads do indeed differ in the amount of accidents happening on them. Finally, Evans showed that even the evidence originally cited to support RHT suffers from methodological shortcomings and appears to refute RHT upon further investigation. There are several other studies taking a similar position on RHT in general (McKenna, 1987, O'Neill & Williams, 1998, Wilde, Robertson & Pless, 2002).

Additional criticism expressed on multiple occasions (Hoyes & Glendon, 1993, Glendon *et al.*, 1996, Adams, 1988) concerns the inability of RHT to be falsified. It is a requirement for any scientific theory to make claims that can be proven wrong and the above authors agree on the view that RHT is, for all practical purposes, not falsifiable. Hoyes and Glendon state that RHT can theoretically not be falsified (as falsification itself is at its core a theoretical concept), while Adams describes RHT as plausible, but untestable. Glendon *et al.* form the proposition that RHT is not falsifiable "in the field", as not all potential pathways through which homeostasis might happen can be controlled for here, but that a sufficiently sophisticated laboratory experiment might be able to detect the mechanisms through which homeostasis takes place. The view that RHT is not falsifiable in real traffic situations is also shared by Hoyes and Glendon (1993), who add the following point: if the only way to change accident rates is to alter the target level of risk, then how can we tell if the target level of risk has actually changed? Looking at accident statistics as the only measurement for this is circular and therefore fruitless. Additionally, the authors state that if RHT really aims to explain all risk behavior, a drop in traffic accidents might very well be accompanied by a rise in accidents in an unrelated field that will remain undetected. Regardless of the (at times quite heated) debate concerning RHT, maybe the state and place in science of this theory is best summed up by Adams (1988), who described RHT as a metaphysical concept that accounts for behavior which no one has yet succeeded in pinning down and measuring with a tool accepted by scientists. Even though RHT is untestable in the real-world situations it was originally invented for, it is full of insight and meaning that might very well have consequences for how we view risk behavior as a general concept. Considering all these notions, the obvious step is to investigate RHT in the laboratory, where the highest amount of control over the circumstances or potential risk behavior can be exerted.

When testing RHT in the laboratory, there are a few potential pitfalls to be taken into consideration. According to Glendon *et al.* (1996), testing risk compensation or homeostasis in the laboratory greatly shortens the time in which multiple accidents or near-misses happen (for example in the case of a driving simulator), which is why one can expect the homeostatic process to take place within a relatively short timeframe. However, it is absolutely vital to encourage participants to take the experiment very seriously, as of course a "crash" in a simulation does never bear any real danger to the utilizing person. This, if treating the experiment like a real situation is not stressed enough, might lead to participants taking either a lot more risk than usual because there are no real consequences or to take no risk at all, because the associated psychophysiological arousal does not take place as easily or simply does not appear. A general concern that was raised about RHT in laboratory settings was that, because the possibility of real, actual harm is always absent, RHT studies might only become investigations of optimization behavior instead of homeostatic processes (Hoyes & Glendon, 1993). However, authors generally agree (Hoyes *et al.*, 1996, Glendon *et al.*, 1996, Hoyes & Glendon, 1993) that investigating RHT in simulated laboratory settings is the only truly fruitful way to uncover the mechanisms through which risk compensation and homeostasis take place because of the level of control they provide in contrast to field experiments or observations.

1.2 The present experiment

For the present thesis, an experiment was conducted with the aim to uncover possible influences of certain types of messages on risk behavior and with the additional goal to confirm the processes of risk compensation and homeostasis proposed by Wilde (1982a). The idea for the effect of certain types of messages on risk behavior was at first more of anecdotal than of scientific nature. It had been observed that drivers participating in so-called *driver safety training* sometimes altered their behavior in traffic in a way that was not intended by the training. Driver safety training mainly consists of bringing drivers into situations that are objectively dangerous. On a specialized training course, drivers learn to remain in control over their car in these situations as much as possible. As an example, drivers are instructed to drive at a certain speed and after a few meters, a plate under the back wheels is quickly moved to one side, making the car spin around its vertical axis. The safety instructor now teaches the driver how to regain control over the spinning car by utilizing certain steering patterns. This one and other situations serve the purpose of preparing drivers for dangerous situations in traffic and equipping them with the knowledge and skills for reacting correctly in those situations. Additionally, driver safety training also teaches how to assess the signs of a dangerous situation coming up and how to avoid it if possible, *along with the important notion that all dangerous situations should be avoided, if possible*. While certainly a promising procedure on paper, practical observations have shown that drivers who have participated in driver safety training exhibit one of two behaviors: they either become more careful and defensive in traffic as intended, or they drive more recklessly because they now feel like they can get any dangerous situation under control before any harmful consequences occur. While at first an anecdotal observation, it turned out to be indeed scientifically confirmed that driver safety training is far less effective in traffic casualty prevention than assumed (for an overview, see Christie, 2001). Based on this, it was hypothesized that driver training might be counterproductive for some people because instead of stressing what could be *lost* in case of a traffic accident, it stresses (not intentionally, of course) what could be *gained* by using the new knowledge and skills to avoid dangerous situations, leading to people *overestimating* their newfound abilities and *compensating* for the new perceived safety level by engaging in more dangerous traffic behavior. It has therefore further been hypothesized that risk behavior could be influenced by how the message of the effects of behavior is conveyed: it might very well be that positive, gain-framed messages have a different effect on behavior than negative, loss-framed messages do. Research confirms different effects of both kinds of messages on risk behavior (Hwang *et al.*, 2012); however, it was hypothesized that it should also make a difference *how* these messages are conveyed, for example if people are merely educated about the potential consequences of a behavior or if the consequences are actively demonstrated to them (as it is the case in driver safety training). The idea that conveying a message of behavioral consequences either through text or through demonstration (for example by video) might lead to different behavior came to mind because reading about and viewing these consequences employs different mental pathways. While viewing is more easily applicable to the actual situation as it shows the actual situation, it also requires less cognitive focus to watch than it does to read. Watching a video of behavioral consequence is more passive than reading a text about it as reading is one of the most complex actions the human brain is capable of and requires considerable concentration. Two mental pathways differing in their depths of information processing have been proposed on multiple occasions in cognitive psychology, with probably the most common being the Dual Process Theory (for a summarizing overview, see Kahneman (2003)). According to Dual Process Theory, any mental operation and cognitive analysis of reality can be carried out either through the fast-thinking, heuristics-employing *system 1*, or through the reflective, slower, reason-bound *system 2*. System 1 is responsible for fast reactions and intuitive judgments, while system 2 leads to thought-out, educated decisions. Applied to the present experiment, it can be hypothesized that the inherently more passive nature of watching a video can lead to it being processed quickly and without much cognitive impact through system 1, while the requirement of concentration for reading a text causes it to be processes more thoroughly through system 2. However, as said before, the video demonstration of consequences applies better to actual reality than a text does, therefore it could also be expected that the video message is analyzed through system 2 as well. To account for both possibilities, the hypotheses for this experiment did not specify an expected direction of difference between these two conditions. To investigate the different effects of these messages on risk behavior (and to find out if there are any), a video game based experiment was conducted that allowed for precise data gathering.

As the second objective of this experiment was to investigate the risk compensation and homeostasis mechanisms proposed by Wilde (1982a), the video game was programmed to allow for multiple gameplay rounds of varying safety for the player. It is hypothesized that, according to Wilde's theory, the introduction of "barriers" that make the game safer (such as seatbelts in cars, which make driving objectively safer) will not affect the players' desire to be safe and will as such lead to an adaption in risk behavior to compensate for the added safety. In other words, it is expected that the safer a gameplay round is, the riskier the players will behave in the round. How exactly these different levels of safety were implemented is explained in chapter 2.2.

To summarize the present research, an experiment was conducted which employed a video game to answer two main questions. The first question was whether the framing of a message that informed about the consequences of behavior (gain-framed vs. loss-framed) and the way the message was conveyed (information transfer through text vs. visualization of consequences through video) had differing effects on risk behavior. The second question was whether making a round of gameplay objectively safer causes players to behave more riskily, cancelling out the added safety effect as proposed by Risk Homeostasis Theory (Wilde, 1982a).

2 Method

2.1 Participants

The total number of participants was 178, with 129 of them being female, 44 being male and 5 being of unknown gender due to an unfortunate processing error in the questionnaires used to assess gender. Their age ranged from 18 to 57 years (*M*=22.4, *SD*=4.7). All participants were either paid for participation or received course credit.

2.2 The video game and its parameters

The free software GameMaker 8.1 has been employed to program the game used for this study. GameMaker 8.1 features a graphical user-interface designed to aid programming. The game itself was programmed to be run in full screen and the actual gameplay was displayed in a box with a resolution of 800x600 pixels (width by height). When the used screen did not fit this resolution, the unused pixels were filled with a black border around the gameplay part of the screen.

Named SpaceGame, the game featured a 2-dimensional environment with a sprite of a small, white spaceship being the controllable element. The objective of the game was to steer the ship through an array of incoming meteors, which were also displayed as sprites. The sprite for the ship had a size of 90x40 (width by height) pixels and the sprites for the meteors had a size of 60x60 pixels. Both the spaceship and the meteors had rectangular collision boxes that fitted their sprite size. The game allowed the ship to be controlled with the arrow keys of the keyboard on the y-axis (in the upwards and downwards directions). The axis the ship was moving on was positioned 109 pixels away from the left side of the gameplay window. During gameplay, the background image (a schematically displayed city during nighttime) was animated to scroll from the right to the left border of the gameplay screen, creating the illusion of flight from the perspective of the player. The meteors were generated at random points on the y-axis at the right side of the gameplay screen and moved towards the ship displayed on the left side. Once generated, the individual meteors never changed their elevation and only moved on the x-axis (the horizontal plane). The game featured 13 difficulty levels, with each difficulty level corresponding to a set speed of the meteors. The standard speed of the meteors (corresponding to a difficulty setting of 1) was 320 pixels per second. Using the right arrow key, the player was able to accelerate the ship[1] in 13 difficulty steps of +50 pixels per second up to 920 pixels per second. Holding the right arrow key down increased the difficulty level by 1 every 250 milliseconds. Therefore, holding down the key served as a means to quickly accelerate while pressing it briefly allowed the player to accelerate only slightly, providing easy control over the speed chosen. Complementarily, the left arrow key could be used to slow the ship down in the same intervals. In the top-left corner of the gameplay screen, icons of shields were displayed. Each time the ship collided with a meteor, the rightmost shield was depleted and removed from the display. When a collision occurred while there were no shields left, the ship was destroyed, which was indicated by an animation showing the ship exploding. When this happened, the respective round was over. Images of all sprites and an example screenshot of gameplay are found in the appendix.

While the game was being played, it traced certain gameplay parameters and wrote them into a logfile 100 times per second. The parameters logged were: a) the current difficulty level (which could be computed into speed), b) the position of the ship on the y-axis, c) the position on both axes of the leftmost meteor sharing any part of its vertical collision box with the vertical collision box of the ship and d) the position of the meteor closest to the ship on both axes. Parameter c) was logged because it could be used to calculate the time that was left until the ship would collide with the meteor currently in its path when it was not steered out of the way. Parameter d) was logged because it allowed the calculation of the distance between the ship and meteor closest to it. Figure 2 illustrates the parameters schematically. Parameter a) is visibly absent from the illustration as it was not displayed in the gameplay window but kept track of by the game as an internal variable.

[1] In reality, the meteors accelerated, but as the ship was the controllable element, speeding up the meteors created a very believable illusion of a faster flying speed from the ship’s perspective.

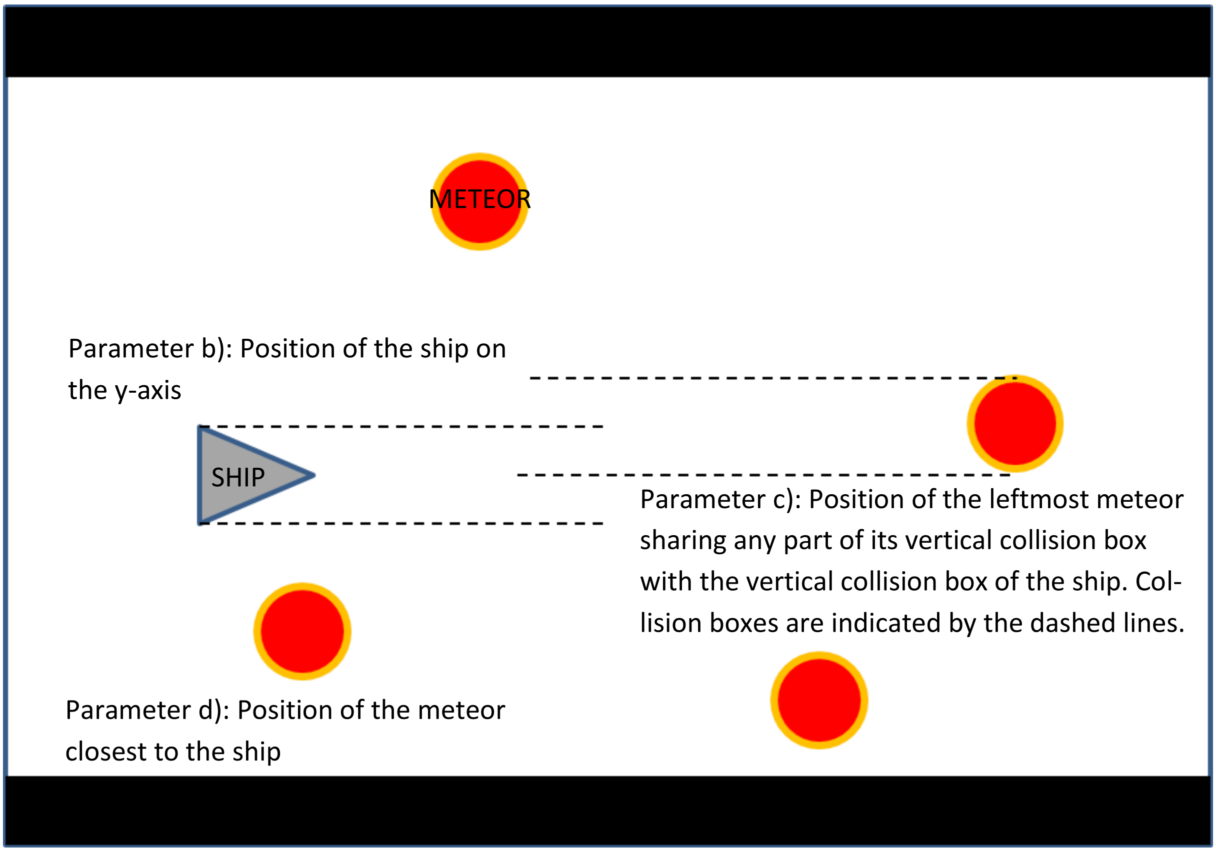


Figure 2: A schematic illustration of the gameplay screen, with written explanations of the parameters that were written into the logfiles. An actual gameplay screenshot can be found in the appendix.

In addition to these parameters, the logfiles also contained the time the respective round took, the number of shields currently active and a unique ID corresponding to the currently played round. The logfiles were saved on the hard drive of the computer the game was run on in the format .csv, allowing for easy usage with Microsoft Excel and SPSS. Once a gameplay session was over, the logfiles could be read by Excel as a table displaying all the aforementioned parameters in a line, while each line represented a single writing process, meaning that there were 100 lines per second of gameplay with each line containing one entry for each parameter. These files were then named steplog.csv and the game placed them in a folder named after a number which could be entered in a field at the start of the game.

The steplog data were used to compute the variables that were later used for analysis. Variables that indicated how much risk participants were willing to take were needed and three different ones were decided on. The first variable was speed, which is intuitively related to risk behavior as people choosing higher speeds take more risk of collision – higher speeds make it more difficult to react to an incoming meteor in time. The second variable was the time that would pass until the collision the ship and a meteor in its path would take place if the ship was not steered out of the way. This variable was labeled *time to collision (TTC)* and was assumed to be a measurement for risk behavior as well: since it was always calculated with the meteor on collision course with the ship, it served as an indicator for how closely participants would let the meteors get to the ship. In this way, a shorter average TTC would indicate more willingness to risk a collision. The third and final variable was the *distance to the closest meteor*, which served as an indicator for how closely participants would let the meteor fly by the ship. Smaller values indicated meteors coming closer to the ship and therefore hinted at the participant taking more risk in their dodging movements. The formulae used to compute these variables are displayed below.

The formula used to compute *speed* was:

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

With *difficulty* always being a value between 1 and 13 found in the logfiles, this formula yielded *speed* values between 320 and 920 pixels per second.

The formula used to compute *TTC* was:

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

The denominator of this formula required speed to already be calculated. This formula is derived from the physical formula used to calculate speed, which has been converted to yield *TTC* instead. The value 109 has to be subtracted from the x-coordinate of the meteor in path of the ship as the ship is not displayed at the leftmost side of the screen but 109 pixels to the right from it (in other words, if 109 was not subtracted in the nominator, the formula would instead yield the time it would take a meteor to collide with the left side of the gameplay screen, at which point it just disappeared). As it is shown here, the formula yielded the current time, in seconds, until a collision between the ship and the meteor in its path would happen.

The formula used to compute the distance to the closest meteor was:

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

The Pythagorean Theorem was used to create this formula, as the ship and the meteor closest to it always formed a triangle with its hypotenuse being the distance between the two. As the game measured all distances and positions in pixels, this formula yielded the distance between the ship and the meteor closest to it in pixels as well.

For the experimental manipulation, texts and videos were created. Participants were separated into 5 groups. One group served as a text-based control group, two of the groups were presented with detailed text-based instructions before playing the game and the two final groups were presented with video-based instructions. The titles of the text- and video-files were never visible to participants during the experiment. The text for the control groups only contained very basic information about the key configurations and controls of the game. The two texts for the next two groups contained the same information about the controls of the game; however, they differed in their approach to convey the gameplay objective. Text A, titled AVOID_COLLISION, stressed the need to avoid a collision of the ship with a meteor at all costs, describing the consequences of a collision with strong words (example: "[…] your ship will be destroyed!"). This text served as a loss-framed instruction. Text B, titled SMOOTH_RIDE, stressed the need for the ship to have a safe flight without mentioning collisions or losses and also conveyed the messages that the shields should not be relied upon and that focus should instead be on the ship staying out of danger. This text served as a gain-framed instruction. The groups receiving video-based instructions were presented with short texts as well, which were read right before the respective instruction-video was played. Video A was described before playback as showing consequences of careless and risky flying and what to avoid during gameplay. The video itself showed the ship starting with 3 shields, accelerating quickly to maximum speed right from the start and repeatedly crashing into meteors either in its path or adjacent to it while dodging, ultimately exploding on the final collision. This video served as a loss-framed instruction, highlighting the potential costs of risky behavior. Video B was described before playback as showing desirable flying behavior. It displayed the ship starting with 3 shields, accelerating slightly (as to not mask the function of accelerating completely from the participants) and dodging all incoming meteors as widely as possible. The ship was shown to slow down when there were too many meteors in close proximity to it and the video ended after 30 seconds with the ship never colliding with a meteor, flying relatively slowly for the entire duration. This video served as a gain-framed instruction, highlighting the potential benefits of careful behavior. All five manipulations were hypothesized to have a different effect on the variables associated with risk behavior, with the hypotheses being two-tailed. No assumptions were made about which method of conveying the messages would results in a larger effect, as explained in chapter 1.2.

2.3 Procedure

Participants were tested on three consecutive days. Testing was done in a room with multiple computers so up to 12 participants could be tested at the same time. The participants were always spaced apart so that they could not see their neighbors' screens. The game was run on each computer but the main files that the game read from and wrote to were saved on a central server instance. Between each of the testing periods, a post-it note with a unique participant number was visibly attached to the screen of each computer. Once all participants for the testing period had taken their seats, one of the experimenters read a short text about the process of the experiment. In this text, it was explained that the objective of the game was to dodge the meteors and "deliver very valuable cargo". After hearing this information, participants completed the first half of a questionnaire (which was part of a different research), proceeded to play the game and completed the second half of the questionnaire. When a participant started the game up, the game called upon a pre-made, randomized stack file to determine which of the 5 instructions (control, Text A or B, Video A or B) should be displayed for this participant. The stack file contained 300 lines, each line with numbers corresponding to one of the instructions. As the stack file was saved on a central server instance, each time a new participant started the game (on any of the computers) it was able to read the topmost unused line and marked it as used afterwards. The numbers in the stack file were randomized with the aid of random.org.

The gameplay itself always started with a dialog asking participants to fill in their participant number as found on the post-it note into a field within the gameplay screen and click a button labeled "OK" afterwards. One of the 5 instructions was then displayed with a button at the bottom center of the screen labeled "start game". The instructions always ended with the cue that a practice round will precede the real gameplay. Once participants read the instruction texts or watched the instruction videos, they pressed the "start game" button. This started a practice round, in which participants started with either 1 or 3 shields, with the number of shields once again being randomly selected from the stack file. The maximum duration of the practice round was 4 minutes. If the ship had lost all shields and was destroyed, the round ended prematurely. When participants were able to not destroy the ship for 4 minutes of gameplay, an animation of the ship flying into background was displayed and the round ended. After the practice round, participants played 5 critical rounds of gameplay, with each round lasting a maximum of 4 minutes. The same animation as in the practice round was played upon completion of a round. At the start of each of the 5 actual rounds, the ship was equipped with 0, 1, 3, 5 or an (to the participants) unknown amount of shields. Each of these conditions was played once and their order was once again randomized with the aid of the stack file. In case the amount of shields was unknown to the participant, a white question mark was displayed instead of the shield icons in the top left corner of the gameplay screen. In actuality, the number of shields in this condition was always 3. Once all gameplay rounds were completed either through lasting 4 minutes each or the ship being destroyed, the game displayed a text telling participants to click the "end game" button, displayed below the text. Once this button was clicked, the game stopped and the second part of the questionnaire opened for the participants. Once this was finished, the experiment ended and participants received their reward for participation. Once all 178 participants had been probed, all folders with logfiles were saved onto a hard drive for further analysis.

The logfiles created by the game, the stack file used for randomization and the formulae described above were used to first create new files for each participant. The new files contained the participant number, the instruction received, the amount of shields in the practice condition, the amount of shields each critical gameplay round started with in the order they were played in and the time each round including the practice round took overall. Also included was the mean speed participants chose for each shield condition and for each amount of shields within each condition. As an example, the files contained the mean speed that a participant chose between when they started with 3 shields and lost one shield, the mean speed for when they then continued with 2 shields until they lost one again, and so on. The same was calculated for the speed modus, the minimum speed and the maximum speed. TTC was treated much the same way, being calculated for each amount of shields for each condition as the mean, median, minimum and maximum TTC. Lastly, the distance to the closest meteor was also calculated in this format as mean and minimum. Once such a file had been created for each participant, these files were then merged into one big file in which the means, minimum and maximum for speed, TTC and distance to the closest meteor were calculated for each participant first for each full round (as opposed to each amount of shields in each round) and then across the entire game. This file, containing all information for all participants, was used for all further analyses.

2.4 Design

There were three independent variables in this experiment. Which instruction participants had received prior to the experiment was manipulated between subjects in 5 steps, while the number of shields a round started with was manipulated within subject in 5 steps. Additionally, the number of shields currently left was used as another independent variable for a second analysis and was manipulated in 17 steps. Used as dependent variables were the mean and maximum speed, mean TTC and mean and minimum distance to the closest meteor. To detect possible homeostatic processes, the time each round took was used as a dependent variable as well.

Possible interactions were accounted for by combining the first two independent variables in a single analysis. To avoid a cumulative type-I error, a mixed-factorial ANOVA was used for all comparisons. To gain additional insight into how the number of shields currently left affects risk behavior, the aforementioned dependent variables (excluding the time each round took) were once again analyzed in a within-subject ANOVA with the amount of shields currently left as the independent variable. For all analyses, partial η^2 is reported as an effect size measure where applicable.

3 Results

In the illustrations in figures 3 to 8 in this paragraph, the different instructions are displayed as numbers from 1 to 5 to allow for clearer legibility in the graphs. In the following table, it is displayed which number corresponds to which instruction.

Number	Instruction
1	Control
2	Text AVOID_COLLISION
3	Text SMOOTH_RIDE
4	Video AVOID_COLLISION
5	Video SMOOTH_RIDE

The relation between which instruction participants received, how many shields a round started with and the speed they chose is displayed in figures 3 and 4. Descriptively, it is obvious that, depending on how many shields the ship was equipped with at the start of a round, participants chose a different speed, as the different curves do not lie on top of each other.

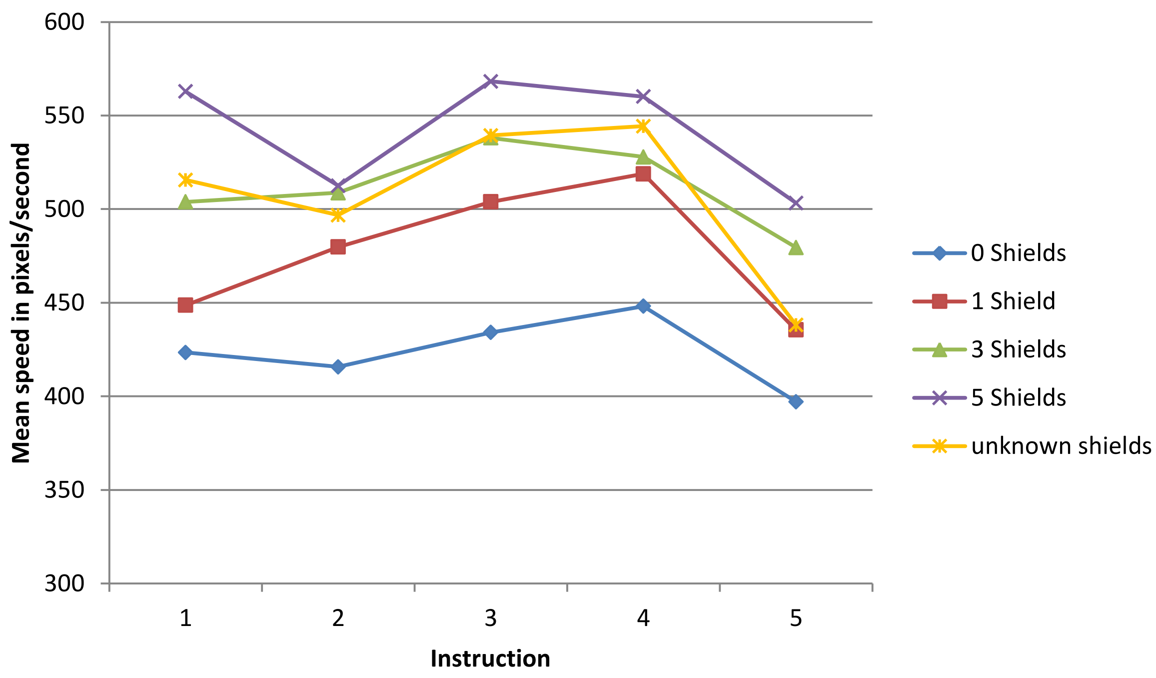


Figure 3: Illustration of the average mean speed chosen by participants as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

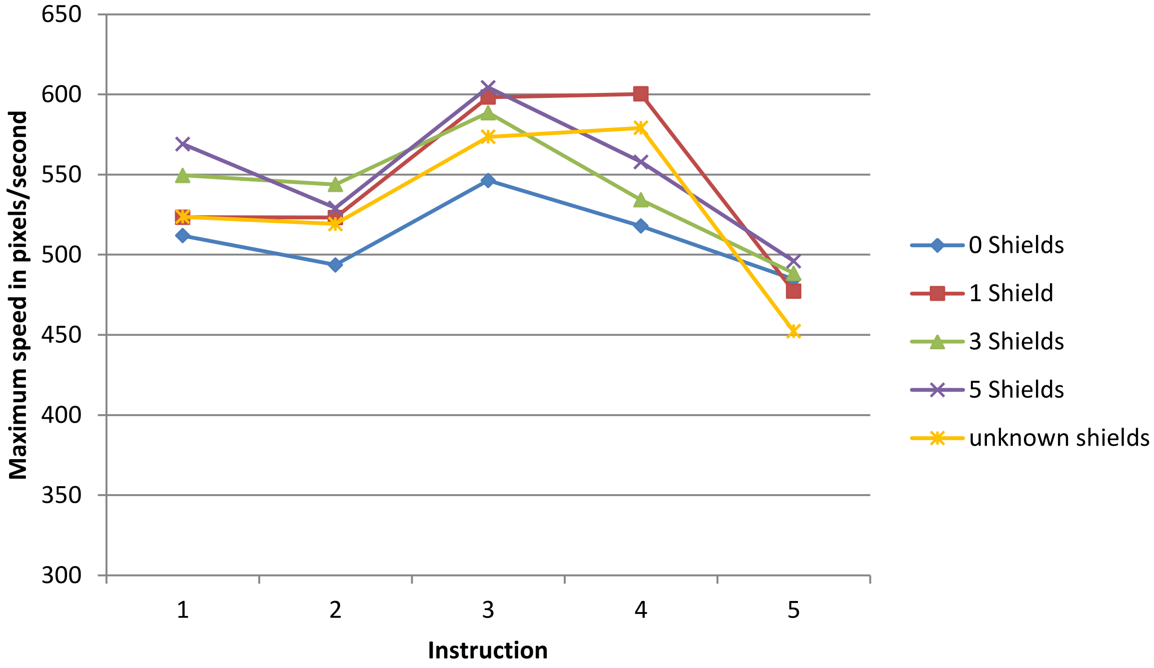


Figure 4: Illustration of the average maximum speed chosen by participants as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

The mixed-factorial ANOVA revealed that the effect of the number of shields on the chosen mean speed was significant, $F(4,169)=52.63$, $p<.001$, $\eta^2=.556$, indicating that mean speed differed significantly depending on how many shields a round started with. Post-hoc Bonferroni-corrected pairwise comparisons revealed that this difference was significant between all groups (all $p<.05$) *except* for the groups “3 Shields” and “unknown shields” ($p=1$). The effect of the instructions on the mean speed turned out to be non-significant, $F(4,172)=2.04$, *n.s*, indicating that mean speed was not significantly affected by which instruction participants had received. Additionally, the interaction between instruction and number of shields for mean speed also turned out to be non-significant, $F(16,688)=1.3$, *n.s.*.

The data showed a comparable trend for maximum speed. The effect of the number of shields on the chosen maximum speed was significant, $F(4,169)=2.93$, $p=.022$, $\eta^2=.065$, indicating that maximum speed differed significantly depending on the number of shields a round started with. Post-hoc Bonferroni-corrected pairwise comparisons revealed that this difference was only significant between the groups “0 Shields” and “1 Shield” (p=.038). Again, the effect of the instructions on the maximum speed was non-significant, $F(4,172)=2.103$, *n.s.*, so maximum speed was not affected by which instruction participants had received. The interaction between instruction and number of shields also was not significant, $F(16,688)=1.03$, *n.s.*.

In figure 5, the relation between which instruction participants received, how many shields a round started with and TTC is displayed. As with speed, the curves differ visibly depending on how many shields the ship had.

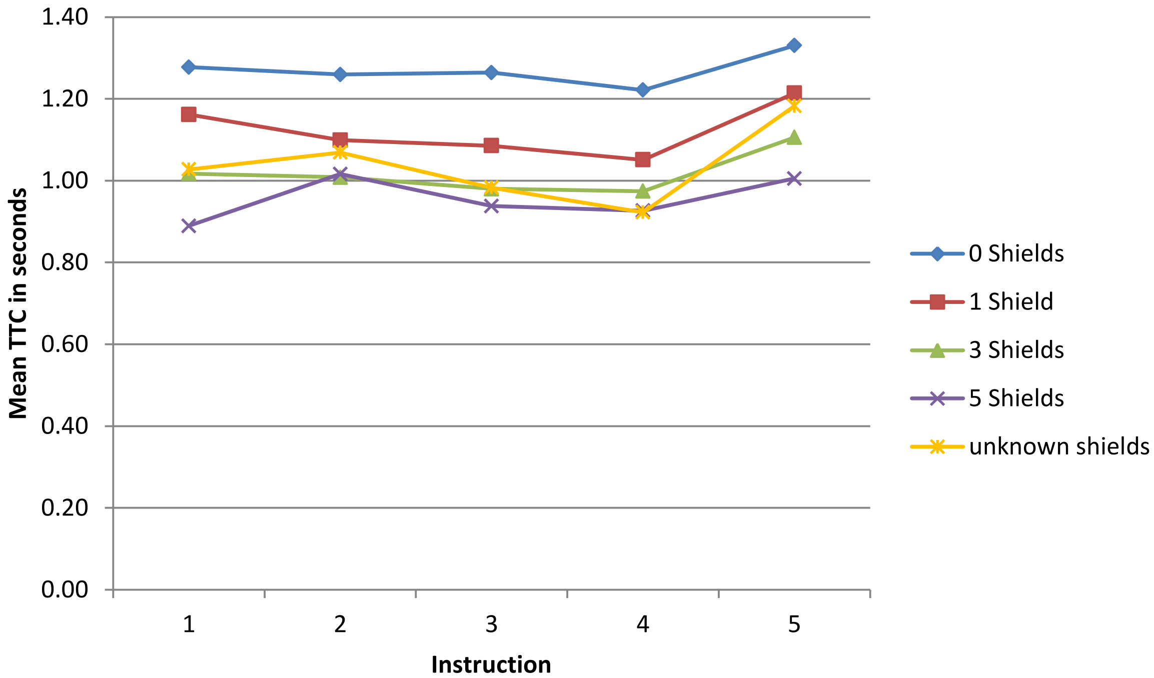


Figure 5: Illustration of the average mean TTC participants had as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

The data for mean TTC exhibited a similar trend as the data for mean speed did. The mixed-factorial ANOVA revealed a significant effect of the number of shields on mean TTC, $F(4,169)=101.327$, $p<.001$, $\eta^2=.706$. Therefore, the number of shields a round started with significantly influenced the mean TTC. Post-hoc Bonferroni-corrected comparisons again revealed that this difference was significant between all groups (all $p<.05$) *except* for the groups “3 Shields” and “unknown shields”. Once again, the effect of the instructions on the mean TTC turned out to be non-significant, $F(4,172)=2.085$, *n.s.*, rendering this manipulation ineffective here as well. The same not significant result was found for the interaction between instruction and number of shields, $F(16,688)=1.625$, *n.s.*.

In figures 6 and 7, the mean and minimum distances to the closest meteor are displayed in relation to the number of shields a round started with and which instruction participants had received. Although the curves are closer together for this measurement, they still differ visibly.

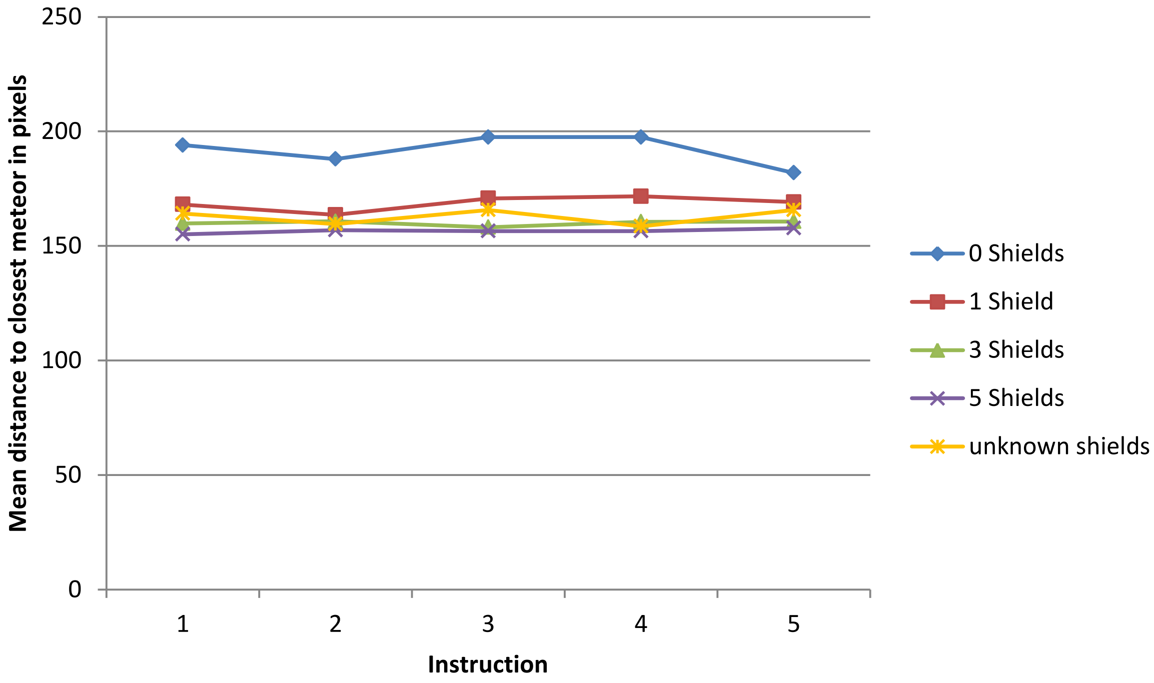


Figure 6: Illustration of the average mean distance to the closest meteor participants had as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

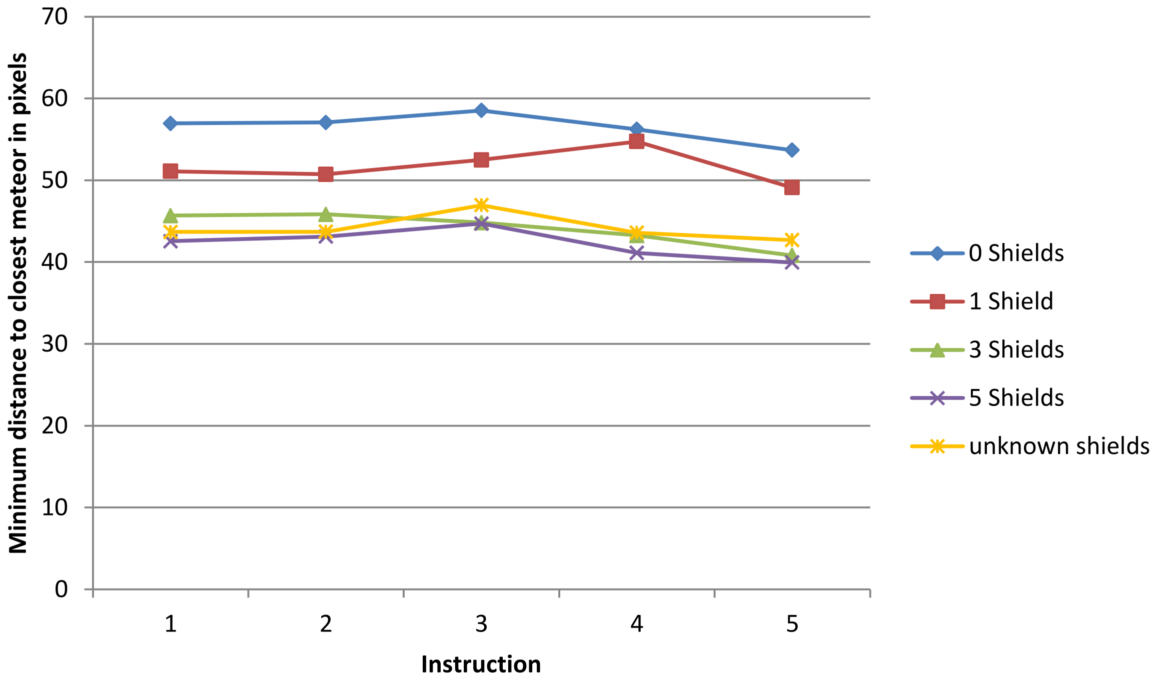


Figure 7: Illustration of the average minimum distance to the closest meteor participants had as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

For the results concerning the distance to the closest meteor, the mixed-factorial ANOVA revealed that the effect of the number of shields on the mean distance to the closest meteor was significant, $F(4,169)=35.351$, $p<.001$, $\eta^2=.456$, indicating a significant influence of the number of shields a round started with on mean distance to the closest meteor. Post-hoc, Bonferroni-corrected comparisons revealed the same trend as for mean speed and mean TTC, with all groups differing significantly (all $p<.05$) *except* for the groups "3 Shields" and "unknown shields" ($p=.509$). Once again, the effect of the instructions on mean distance to the closest meteor was not significant, $F(4,172)=.0613$, *n.s.* and the same was found for the interaction between instructions and number of shields on distance to the closest meteor, $F(16,688)=.963$, *n.s.*. For the minimum distance to the closest meteor, the effect of the number of shields a round started with was also significant, $F(4,169)=39.702$, $p<.001$, $\eta^2=.483$. Post-hoc, Bonferroni-corrected comparisons revealed significant differences between the group "0 Shields" and all other groups (all $p<.01$), significant differences between the group "1 Shield" and all other groups (all $p<.01$) and no significant differences between the groups "3 Shields", "5 Shields" and "unknown shields" (all $p>.05$). The effect of the instructions on minimum distance to the closest meteor was not significant, $F(4,172)=1.745$, *n.s.* and the same result was yielded for the interaction between instructions and number of shields on minimum distance to the closest meteor, $F(16,688)=.541$, *n.s.*.

The relation between the time each round took and the number of shields each round started with and which instruction participants had received is displayed in figure 8.

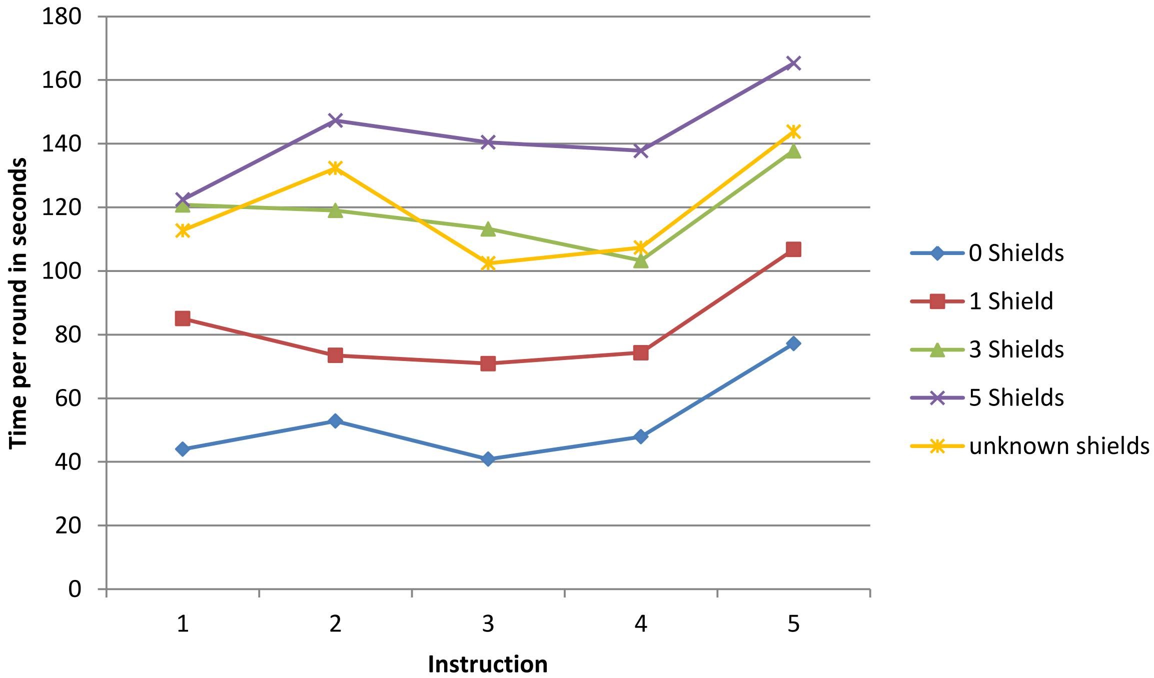


Figure 8: Illustration of the average time each round took as a function of which instruction was received. Differently colored graphs are used for the different amount of shields the rounds started with.

For this measurement, the mixed-factorial ANOVA revealed a significant effect of the number of shields a round started with on time, $F(4,169)=72.116$, $p<.001$, $\eta^2=.629$. Post-hoc, Bonferroni-corrected comparisons showed that this difference was significant between all groups (all $p<.001$) *except* between the groups “3 Shields” and “unknown shields” ($p=1$). When taking the graphs in figure 8 into consideration, these results indicate that the more shields a round started with, the longer it took, which should be inferentially expected. The effect of the instructions on the time once again turned out to be non-significant, $F(4,172)=1.917$, *n.s.*, as did the interaction between instructions and number of shields, $F(16,688)=.713$, *n.s.*.

The next figures display the results of the within-subject analysis that concerns the number of shields left at any point. In figure 9, the relation between the current number of shields and average mean speed is displayed.

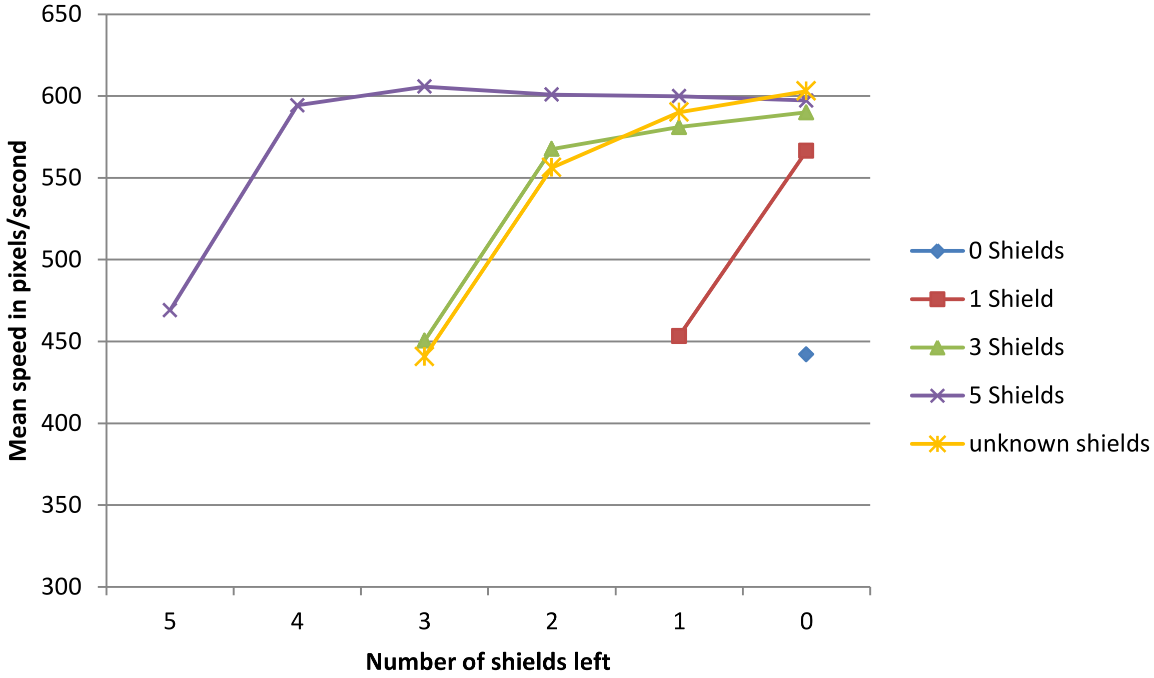


Figure 9: Illustration of the average mean speed as a function of how many shields the participants had left. Differently colored graphs are used for the different amount of shields the rounds started with.

Descriptively, it is visible that the start of each round (when no shields were lost yet) differs from the rest of the respective round along with the non-starting shields of the other rounds. In other words, all first shields of all rounds seem to form a group that differs from the other shields, which also seem to form a group. The statistical comparisons confirm just that, with one outlier that warrants further discussion. The within-subject ANOVA revealed a significant effect of the amount of shields left on mean speed, $F(16,2832)=50.731$, $p<.001$, $\eta^2=.306$. Bonferroni-corrected post-hoc comparisons revealed the trend just described: The first shield of each round differed significantly from both the other shields of its own round as well as the non-starting shields of the other rounds (all $p<.05$). When taking figure 9 into consideration, this indicates that mean speed was significantly *lower* within the first shield of each round compared to the rest of the round. No significant differences emerged between the rest of the data points once one shields was depleted (all $p>.05$), with an exception found in the “unknown shields” round: When participants had two shields left in this round (after their first collision), their speed *increased* significantly compared to when all shields were intact in this round and in the other rounds (all $p<.05$), however, this specific condition (unknown shields, after the first collision) also differed significantly from the rest of the “unknown shields” round (both $p<.05$), meaning that mean speed *increased* even further after a second shield was lost in the “unknown shields” round. Furthermore, the condition in question did not differ significantly from any other condition (all $p>.05$).

Figure 10 shows the relation between the current number of shields and the average maximum speed chosen by participants.

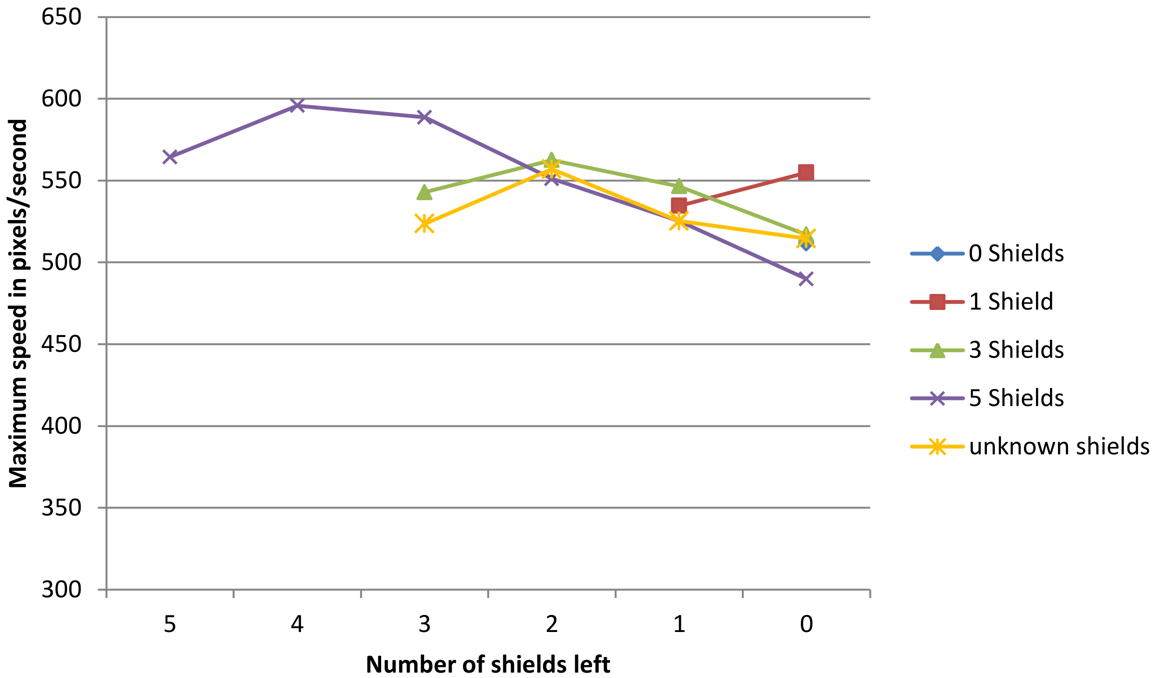


Figure 10: Illustration of the average maximum speed as a function of how many shields the participants had left. Differently colored graphs are used for the different amount of shields the rounds started with.

The data for average maximum speed do not show as clear of a trend as the data for average mean speed. Descriptively, an increase in maximum speed is displayed after the first shield of each round is lost, while maximum speed declines continuously after that point. The within-subject ANOVA revealed a significant effect of how many shields were left on maximum speed, $F(16,2832)=6.005$, $p<.001$, $\eta^2=.034$. The effect size calculated for this result is considered small. Bonferroni-corrected post-hoc comparisons revealed that the differences between the first shield and the second shield in each round were never significant (all $p>.05$), except for the “unknown shields” round, were maximum speed increased significantly after the first shield was lost (*p*=.038). The rest of the values in the "unknown shields" round did not differ significantly from each other (all *p*>.05). For the other rounds, significant differences did not emerge between shields in succession; however, the following differences were found to be significant: the last two shields of the "5 Shields" round differed significantly from both each other and the rest of the shields in this round (all *p*<.05), the "0 Shields" round differed significantly from the first three shields of the "5 Shields" round and from the second shield of the "3 Shields" round (all *p*<.05). The second shield of the "5 Shields" round differed significantly from the first shield of the "3 Shields" round (*p*=.025).

In figure 11, the relation between the amount of shields left and average mean TTC is displayed. Descriptively, it is very similar to figure 9 (mean speed).

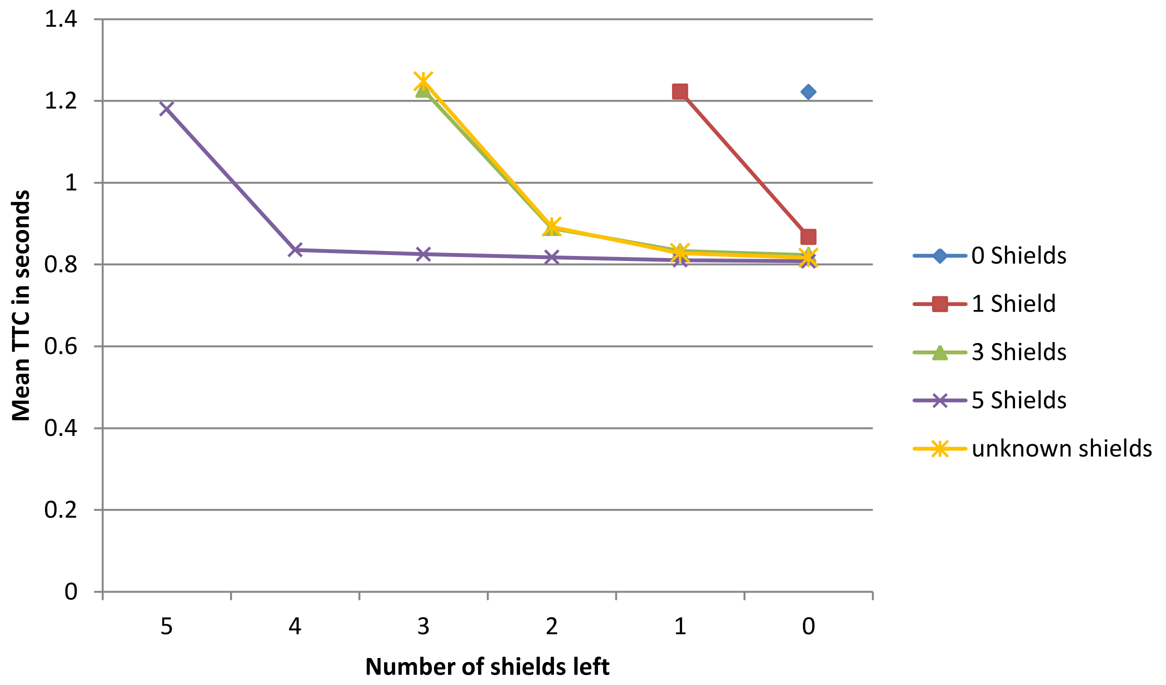


Figure 11: Illustration of the average mean TTC as a function of how many shields the participants had left. Differently colored graphs are used for the different amount of shields the rounds started with.

The data for mean TTC exhibited a trend comparable to those for mean speed. The within-subject ANOVA revealed a significant effect of the number of shields left on mean TTC, $F(16,2832)=75.8$, $p<.001$, $\eta^2=.397$. Bonferroni-corrected post-hoc comparisons confirmed the visual description of the data, with significant differences emerging between the first shield of every round (before a collision) and the rest of the shields both within the same round and the non-starting shields of the other rounds (all $p<.05$). In other words, just like for mean speed, the first shields of each respective round formed a group differing significantly from both the other shields in their respective group and the other shields in all other groups. In contrast to what was found for mean speed, there were no outliers in the data for TTC that differed from this trend.

The relation between the number of shields left and the mean distance to the closest meteor is displayed in figure 12. Again, this variable visibly exhibited the same trend as mean speed and mean TTC.

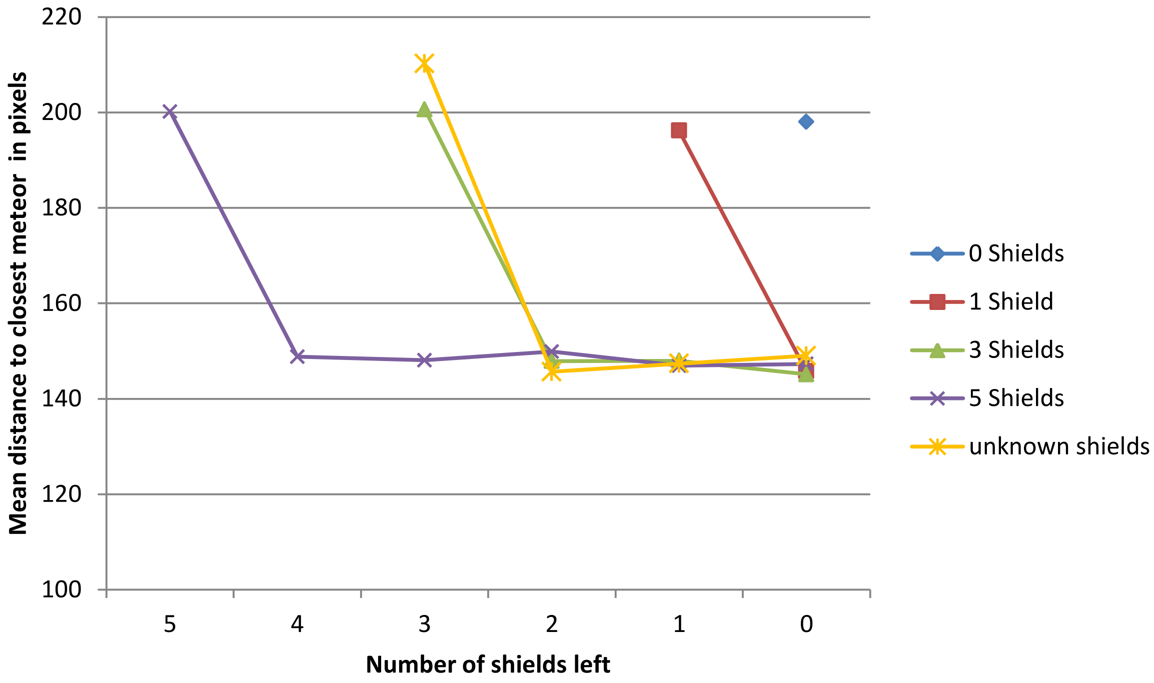


Figure 12: Illustration of the average mean distance to the closest meteor as a function of how many shields the participants had left. Differently colored graphs are used for the different amount of shields the rounds started with.

The results of the within-subject ANOVA once again confirmed the visible impression that the data convey: The number of shields left had a significant effect on the mean distance to the closest meteor, $F(16,2832)=86.593$, $p<.001$, $\eta^2=.43$. Bonferroni-corrected post-hoc comparisons showed significant differences between the first shields of each round compared to the other shields, both of the respective round and the other rounds (all $p<.05$), with all other differences being non-significant (all $p>.05$). Worded differently, the first shields of each round formed a group that differed significantly from the group formed by all other shields.

The next figure shows the relation between the minimum distance to the closest meteor and the number of shields currently left. The trend observed in the data so far is once again visible here, although not quite as clearly. For minimum distance to the closest meteor, the values generally become *smaller* as the round proceeds, indicating that with the progression of the round, participants dodged the meteors more closely than at the beginning of the round.

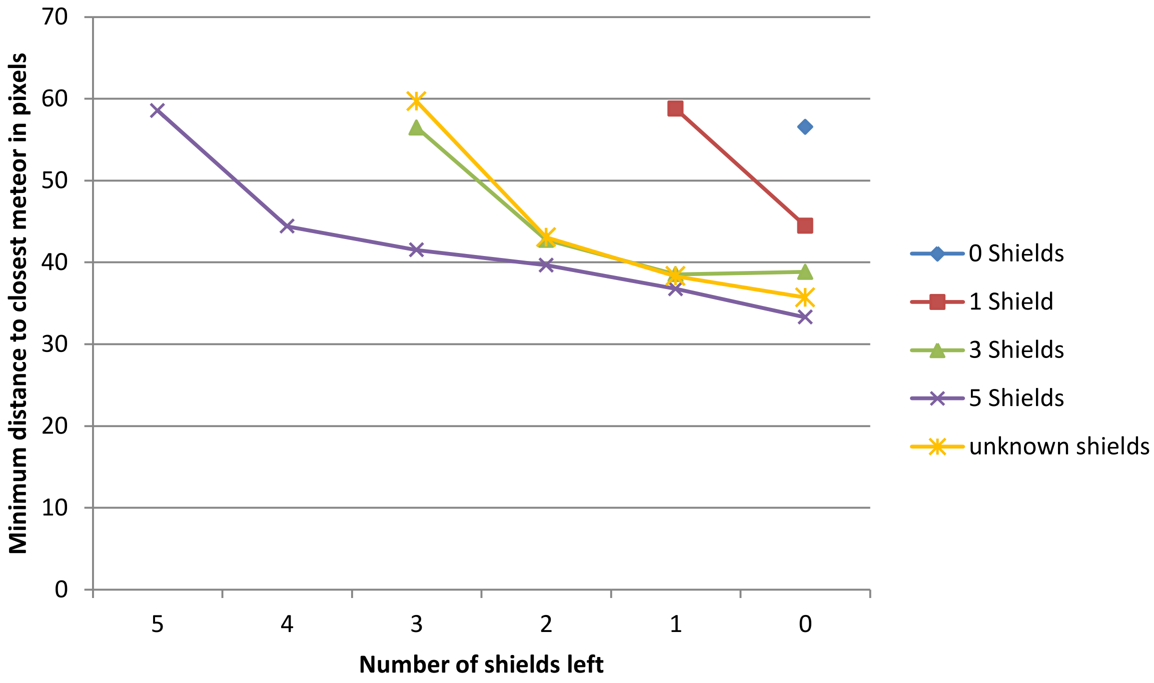


Figure 13: Illustration of the average minimum distance to the closest meteor as a function of how many shields the participants had left. Differently colored graphs are used for the different amount of shields the rounds started with.

The within-subject ANOVA revealed a significant effect of the number of shields left on the average minimum distance to the closest meteor, $F(16,2832)=52.218$, $p<.001$, $\eta^2=.232$. Bonferroni-corrected post-hoc comparisons once again showed significant differences between the first shields of each round and the other shields of both the respective round and the other rounds (all $p<.05$), meaning that the first shields of each round once again formed a group differing significantly from a group formed by all other shields. In addition, significant differences emerged between the second shield of the “5 Shields” round and the second to last and last shield of the same round (both $p<.005$), between the third shield of the “5 Shields” round and the last shield of the same round ($p=.005$) and between the second and the last shield of the “unknown shields” round ($p=.01$).

4 Discussion

The aim of the present experiment was twofold. A video game featuring a spaceship dodging obstacles in flight was programmed to measure and log parameters of risk behavior during gameplay. The first objective was to investigate the influence of gain-framed and loss-framed messages conveying the consequences of behavior on risk behavior. Both types of messages were presented either as a text or as a combination of a text and a video, with the assumption that the text-video combination would differ from the text-only instructions in their effects on risk behavior as the video directly demonstrated the consequences of a certain behavior, while the text merely described it, but required more concentration to actively process. The second objective was to test the mechanisms of risk compensation as proposed by Risk Homeostasis Theory (Wilde, 1982a). This was realized by setting up multiple gameplay rounds with varying amounts of objective safety: through the introduction of shields that protected the spaceship in case it hit an obstacle, having more shields was naturally safer. In line with RHT, it was hypothesized that participants will change their flying behavior to be more risky when the ship is equipped with more shields, ultimately cancelling out the added safety by taking more risk. This is because RHT proposes that people are characterized by a desire to experience a certain degree of safety and that this specific degree is not changed by making actions inherently safer. This was investigated further by analyzing whether or not participants adjusted their behavior depending on how many shields they had left in a round.

The results of the experiment showed support only for one of the two proposed effects. Generally speaking, the idea that gain-framed or loss-framed messages have an effect on risk behavior was not supported. However, the amount of shields (and therefore, the amount of objective safety) in a round did have a significant influence on the measurements used here to assess risk behavior.

The finding that the type of instruction participants had received did not influence risk behavior has some implications. It is imaginable that the way the instructions were conveyed (either through text or through text + video) was simply not strong enough to trigger a response in risk behavior. This might have been due to weaker formulations in the texts or due to the behavior shown in the videos not being shown in extreme enough manners. Nevertheless, the graphs in figures 3, 4, 5 and 8 do show a trend of the two video-based groups (groups 4 and 5) differing in risk behavior and although no *p*-values reached statistical significance, they came very close to in the case of the difference between the two video-based groups. This notion points towards the sample size in the current experience being too small; while the differences between the other groups are questionable, statistically significant differences may have emerged between the two video-based groups with a larger overall sample. Should such differences emerge, however, they should be interpreted with caution. It might be that the group that is shown careful behavior exhibits this behavior not because its value is highlighted but simply because careful behavior was made more salient by showing this video. A comparable mechanism is of course imaginable for the group where the avoidance of risky behavior is highlighted: By showing this behavior, it becomes more salient for participants, therefore they engage in it regardless of being warned not to behave as shown. Such a finding would also mean that showing undesirable behavior together with the instruction that is should not be exhibited is ineffective in suppressing that behavior as it only increases its salience. While the present data do not support these hypotheses quantitatively, they do show a trend pointing into their direction. If future research reveals the findings described here, they should be interpreted with caution and maybe tested for without using gameplay rounds of different objective safety.

Fortunately, the present data did allow for inferences about the influence of added safety on risk behavior. Depending on how many shields they started a round with, participants played the game more riskily, with a higher number of shields leading to higher mean speed, lower TTC values and lower distances to the closest meteor. Although not all dependent variables showed the exact same tendency in differences between the groups, all results point into the same direction.

The differences found in mean speed seem to be very much in line with the notion of risk compensation. The more shields a round started with, the faster participants flew the ship on average, meaning that they compensated for the added safety by behaving in a riskier fashion, with differences being significant between all groups except for two (this finding will be discussed in detail below). The differences found in maximum speed are less pronounced than the ones in mean speed, with significant differences emerging only between two groups. This might be because maximum speed was only exhibited during a few moments of gameplay, likely when participants tried out how fast they can go, probably during the beginning of a round when the number of meteors on the screen has not reached its maximum yet. Overall, maximum speed might not be as valid as a measurement of risk behavior as mean speed is, as mean speed takes the entire data of one gameplay round into consideration. This idea is supported by the fact that for maximum speed, only the groups “0 Shields” and “1 Shield” differed significantly, which are not at opposite ends of the spectrum of the number of shields possible, but quite close to each other. This makes the validity of maximum speed as a measurement of risk behavior in this experiment very questionable. Regardless, the data for mean speed clearly support the notion that people adjust their risk behavior depending on how objectively safe the situation is, which is in line with the mechanics for RHT proposed by Wilde (1982a).

The data for mean TTC in the mixed-factorial ANOVA exhibited a trend comparable to mean speed. Once again, the higher the number of shields was, the lower the mean TTC for the respective round was, indicating that on average, participants let meteors get closer to the ship before dodging out of the way when the round had started with more shields. Just like the finding before, this one also points towards risk compensation, as participants compensated for safer circumstances by putting the ship in more dangerous situations. Even though this finding seems to confirm the first one, it should be interpreted with caution: As the formula used to calculate TTC necessarily needed to include the current speed, TTC changed depending on which speed participants chose, so the two measurements might be interrelated to a large extend. Additionally, TTC was always calculated with the meteor in path of the ship - which meteor that was changed quicker depending on how fast participants were able to react, which was again dependent on speed. Faster speed of course brings the ship closer to the object on collision course than slower speed before a dodging reaction can take place (assuming equal reaction times). Considering these notions, it is imaginable that TTC is, in this case, not a highly valid measurement for risk behavior, as it depends too much on speed in the two ways just described.

For the mean distance to the closest meteor, the findings were once again comparable to mean speed (and mean TTC), as all groups differed except for two. The direction of this difference was once again in line with risk compensation: The higher the number of shields was, the lower the mean distance to the closest meteor was, indicating that participants once again compensated for a safer situation by bringing the ship closer to the objects it could collide with. These differences were, once again, significant between all groups except for two. The findings for minimum distance to the closest meteor were not as clear-cut. Significant differences emerged between the group "0 Shields" and all other groups and the group "1 Shield" and all other groups, with no additional differences found. Taking the graphs in figure 7 into consideration, it is clear that the direction of these differences is in line with risk compensation, as the "0 Shields" group has the lowest minimal distance to the closest meteor while the group "1 Shield" had the second lowest. The fact that no difference emerged between the other groups might indicate that minimum and mean distance to the closest meteor both measure different aspects of risk behavior (or measure different things altogether). Minimum distance to the closest meteor is far more heavily influence by very short moments in time where a single meteor is dodged closely than mean distance to the closest meteor is. Of course, these near-misses are the moments of the most imminent danger, meaning that they might be even more valid measurements for risk taking behavior, as average minimum distance to the closest meteor becomes smaller the more near-misses participants produce. However, having a small number of shields might also increase nervousness and therefore impair concentration, artificially increasing the number of near-misses. As the data do not allow controlling for this potential influence, mean distance to the closest meteor should be viewed as a more valid measurement for risk behavior in the context of this experiment.

Finally, the results concerning the time each round took are both surprising and not. Not surprising is the finding that the more shields a round started with, the longer it took, which is to be expected as more shields mean more buffers that protect against the destruction of the ship, extending the time that the round can be played for before the ship is destroyed. The surprising finding is that this is not in line with RHT: RHT would have predicted equal times for each round irrespective of the number of shields, as shields do not increase the participants' desire to be safe, therefore they should have no effect on overall time of the gameplay rounds. RHT would have predicted a compensation for the added safety that is so complete that each round ends in fatal destruction after the exact same amount of time, irrespective of the number of shields. A complete compensation according to RHT would mean that crashes occur more often with a higher number of shields, keeping the time after which a fatality (the destruction of the ship, which would equal to the death of a driver in real traffic) constant.

Unfortunately, the results of the within-subject ANOVA that had the number of shields currently left as the independent variable seriously challenge the results of the mixed-factorial ANOVA. The results for mean speed, mean TTC and mean distance to the closest meteor all show the same trend: In each round, the first shield differs significantly from the rest of the shields in both the same round and in the other rounds. The first shield in each round also does not differ significantly from the first shields in the other rounds. For mean speed, mean TTC and mean distance to the closest meteor, these findings point towards two different systematic errors in the data gathering method employed here. As mentioned, the ship always started each round with the minimum possible speed of 320 pixels per second. As the logging of the data started right at the beginning of a round and logging happened 100 times per second, the first few *hundreds* of data points had the minimum possible value of speed, before participants started to speed up. Because of the way means are calculated and the fact that they are very vulnerable to outliers, the first few seconds of each round act as a huge bias in the data that only affect the first shield of each round. In other words, the fact that each round started with the minimum possible speed systematically decreased the mean speed and mean TTC for the first shield of each round drastically. The same is true for mean distance to the closest meteor, albeit for a different reason. When each round started, there were no meteors on the screen. The first meteor was generated as soon as the round started (with all subsequent meteors being generated at a fixed time interval, as described in chapter 2.2); however, as new meteors were always generated on the right side of the gameplay screen while the ship was positioned towards the left, the meteor closest to the ship was actually much further away from it in the beginning of each new round than it was during normal gameplay, when the screen was always filled with a set number of meteors. This introduces a systematic bias much like the initial speed of each round does: Because the distance to the closest meteor was much higher than usual at the beginning of each round, it is systematically higher within the first shield of each round as well. These biases are problematic for the mixed-factorial ANOVA results as well. The means for the dependent variables used in the mixed-factorial ANOVA were calculated across on entire round. Because the data for the first shield of each round were affected by the bias described, they also heavily affected the mean data for the entire round and *therefore also affected the differences between the rounds*. As an example, grand means for the "5 Shields" round were not as heavily affected by the first shield as the grand means for the "1 Shield" round were because the "5 Shields" round had much more additional data (through having 4 more shields) to "counteract" or "outweigh" the initial bias of the start of the round. This scheme can of course be applied to the differences between the other rounds as well and unfortunately also explains the fact that no differences emerged between the "3 Shields" and "unknown shields" rounds very well, as they effectively had the same number of shields. This bias seriously calls all conclusions drawn from the results of the mixed-factorial ANOVA into question, making the ones for mean speed, mean TTC and mean distance to the closest meteor essentially non-interpretable; however, the means for maximum speed and minimum distance to the closest meteor cannot have fallen victim to this bias, as the values they produced in the very beginning of each round did not affect the final data. They therefore need to be investigated separately. The next paragraph will list possible explanations for the findings concerning mean speed, mean TTC and mean distance to the closest meteor, accounting for the bias.

Fortunately, as only the first shield of each round is systematically biased, the results of the within-subject ANOVA for the other shields of each round are interpretable in relation to each other. However, it quickly becomes clear that for mean speed, mean TTC and mean distance to the closest meteor, there are no differences between the shields after the first one at all with only one exception for mean speed within the "unknown shields" round. This means that risk compensation is completely absent for these three dependent variables, as participants did not adjust their mean speed, did not show a change in TTC and did not fly further away from the meteors on average when the round became objectively less safe.

However, this does not mean that these data disprove RHT: As there was no risk compensation found for these variables, the participants' desire for safety seemed to be unaffected and, as a matter of fact, seemed to be the only thing dictating their risk behavior. In other words, for mean speed, mean TTC and mean distance to the closest meteor, instead of adjusting their risk behavior accounting for the degree of objective safety, participants seemed to be guided by an "inner safety compass", not taking the situation they found themselves into account and instead relying on their own desire for safety to asses which speed to choose and how closely to dodge the meteors. Wilde (1982a) proposes that peoples' inner desire for safety, in the long run, is the only thing that their risk behavior adapts to, which is seemingly in line with this interpretation of the present results. This interpretation must, however, be viewed with caution, as the concept of a desire for safety in a laboratory setting with the absence of real danger is at least questionable, as pointed out in chapter 1.1 already. The outlier found in the trend of the data within the "unknown shields" round is an interesting additional finding: The second shield in this round differs significantly from the second to last and last shields, albeit only for mean speed and not for mean TTC or mean distance to the closest meteor. This is puzzling, as in the absence of any visible information about the objective safety of a round, participants still adjusted their risk behavior towards a value that is comparable to the value of the last shields of the other rounds. However, this adaption is slower when compared to the other rounds. One possible explanation is that, in the absence of information about objective safety, the starting speed of a round serves as a behavioral "anchor" that participants initially use as an orientation, but once they start losing shields and realize that the "unknown shields" round has more than 0 or 1 shields, the anchor stops having an effect and participants instead once again rely on their own desire to be safe, regardless of the true, objective safety level being invisible to them all throughout the round. This strengthens the assumption that while after some time the objective safety level ceases to matter, with participants' inner desire for safety taking over, but that at least initially, the level of objective safety seems to be taken into consideration as long as it is not visible.

For maximum speed, a general trend that could be observed was that after the first shield in each round was depleted, the average maximum speed increased, only to decrease back down for every additional shield lost during the respective round. Statistically, not all differences in this trend emerged as significant and only the “5 Shields” round showed clearly interpretable results: When left with 4 shields in the “5 Shields” round, participants had a significantly higher maximum speed than when they had only 1 or 0 shields left in the same round. This is the first *unbiased* result that clearly points towards direct risk compensation: As the round continues, it becomes more and more objectively unsafe through the loss of shields and participants adjust their maximum speed to lower levels. This finding contrasts with the ones for mean speed as actual compensation for the objective safety of a round can be found here. The fact that participants showed lower maximum speeds towards the end of the round might indicate a different mechanic of risk behavior: in real-life traffic situations, the moments were maximum speed is driven at are the ones where most accidents can be expected to happen, while driving at average speed is inherently safer. It can be inferred that, in the present experiment, participants increased their own safety not by slowing down on average, but by reducing the maximum speed they were willing to fly at once the level of objective safety in a round decreased and the danger of a “fatal” crash became more and more imminent. As moments of maximum speed are the most dangerous in real-life traffic as well, this finding supports the notion that there is risk compensation in specific areas (such as maximum speed), but not overall.

The data for minimum distance to the closest meteor show a trend that is very puzzling at first. Overall, the minimum distance becomes *smaller* as the round continues, which means that the less objectively safe the round becomes, the *closer* participants came to the meteors in at least one occasion. In other words, during a few short moments for each round, participants took *more* risk the more *unsafe* the round was. Upon further investigation, it was revealed that the entire data for minimum distance to the closest meteor are heavily biased as well. As average minimum distance was only ever calculated with the smallest value found for each shield and then averaged across all participants for this shield, it is very susceptible to numerical outliers as well. In the moment of a collision, the minimum distance to the closest meteor has a value of 0 by definition, meaning that for all events of a collision, the respective participant had 0 as their minimum distance for this shield. What this essentially means is that for every collision that happened to a participant, the average minimum distance for this specific shield was biased more towards zero. The downward trend in the data can be explained as follows: a few participants were always able to finish the rounds with a certain number of shields intact, meaning that they, for example, only lost 2 shields in the “5 Shields” round before the round automatically ended after 4 minutes. For these participants, the minimum distance for the rest of the round was, by definition, a missing value. This means that the less shields were left, the less participants actually came this far down in their number of shields, meaning that each collision that resulted in a distance value of 0 affected the means more the further into a round it happened. To give an example using the “5 Shields” round: there were less participants who had 1 shield left than there were participants with 4 shields left at any point of the round, meaning that a collision and therefore a distance value of 0 had a higher mathematical “weight” for the calculations for 1 shield (left) because these calculations were done with less overall data points than the ones for 4 shields. Because of this bias, minimum distance to the closest meteor should not be used as a measurement for short moments of maximally risky behavior at all in the context of this experiment. A better alternative for future studies would be to use the bottom 5% of distance values for calculations, greatly reducing the impact that values of 0 caused by collisions have on the means.

One highly interesting finding of the present experiment, taking the results of the within-subject ANOVA into consideration specifically, is that for all dependent variables (speed, TTC and distance to the closest meteor), almost no differences emerged between the groups “3 Shields” and “unknown shields”. As mentioned in chapter 2.3, there was no way for the participants to tell how many shields the ship was equipped with in the “unknown shields” round, which is why it was expected that participants would treat this round similar to the “0 Shields” condition – as there was no way to reliably assess the amount of safety in this round, the logical way to play this round would be to fly as slowly and carefully as possible. However, the data show another trend: Even though participants did not know that the amount of shields in the “unknown shields” round was always 3, they still exhibited behavior that did not differ significantly from the behavior they exhibited in the “3 Shields” round. This finding is puzzling, as risk compensation *requires* the amount of perceived risk to be assessed accurately, which was not possible in the “unknown shields” round. While the safest assumption would be to assume 0 shields and no protection, participants did not behave this way. A possible (but very frightening) explanation for this is in line with RHT: even when the objective safety of a situation is unknown, the amount of risk people are willing to take is unaffected, causing them to make assumptions about the situation based on *nothing* and therefore showing a certain level of risk behavior that is much closer to the amount of risk they are willing to take than to the amount of risk they should be taking. People assuming a safety level that is a lot higher than their information allows them to has large implications for road safety: Not knowing the amount of risk (because of, for example, “invisible” risk factors such as very thin ice on the road) might cause people to assume that there is no risk instead of assuming a moderate or maximum amount of risk. The present data seem to point in the direction that when faced with an unknown amount of safety, people seem to treat it like an “average” amount of safety (in the context of this experiment, the mean amount of shields across all rounds was close to 3), which is not as detrimental as assuming absolute safety / no risk, but certainly not optimal. However, there is another possible explanation for this finding that does not shed such a bad light on humans’ ability to assess risk: although people did not know how many shields the ship had, they might have been implicitly able to tell the approximate amount based on the amount of collisions that took place in that gameplay round, making their judgment implicitly more accurate than allowed by the displayed information. From this point of view, it could be inferred the assessment of risk in a given situation is not a conscious process that only takes in visible information, but is instead a more complex, subconscious process that takes all available information into account and therefore is a lot more precise than initially assumed. This notion is supported by the fact that for mean speed, the second shields of the “3 Shields” and “unknown shields” round differed, but towards the end of both rounds, mean speed shifted towards a common value. This might indicate that, the more a round progresses, the more accurate the *implicit* risk assessment becomes. Future experiments could test which of these two assumptions is true by changing the amount of shields that the “unknown shields” round starts with while keeping the other rounds the same. For example, when the amount of shields in the “unknown shields” condition is 5, do the curves for the dependent variables suddenly no longer differ between “5 Shields” and “unknown shields”, or do they not change compared to the present experiment? The former would indicate a mechanic of risk and safety assessment that is implicit and depends on much more than only visible information. The latter would mean that risk assessment in case of no visible information "defaults" to a value that is roughly the same as the mean amount of safety across all rounds. Both explanations would have implications for traffic safety: The first one would indicate that providing as much information (visible and non-visible) as possible about the current risk level is beneficial and aids people in assessing the amount of objective risk correctly, the second one would indicate that information should be given that artificially increases the amount of perceived risk to deter people from assuming a "default" risk. However, if RHT is true in its entirety, then both approaches would yield no changes in fatality rates in the long run as both do not affect drivers' desire to be safe.

To provide further clarity on the mechanics of risk assessment in this type of experiment, the amount of shields in the "unknown shields" condition could be made *unlimited* – is it possible that keeping participants in a constant state of not knowing their safety could reveal a "true" equilibrium of risk behavior that they shift towards. However, if participants are implicitly able to tell that their shields are unlimited (which they could reasonably assume after a large number of collisions), they might just speed up to maximum speed for the remainder of the round without fear of consequences. Again, both findings would have different implications for risk assessment in the context of this experiment. The emergence of an equilibrium (in speed, TTC and distance to the closest meteor) would indicate that there is indeed an internal level of perceived risk that participants are not willing to go over. In contrast, a constant increase in speed up to the maximum would either indicate that participants are (implicitly) able to tell that that their shields are unlimited, or that every collision is viewed as evidence that there are more shields than initially assumed, justifying a speedup. The latter would be highly counterproductive, as even with unknown shields, losing a shield logically means that the risk of destruction is now greater than it was before the shield was lost. Overall, the "unknown shields" round and its comparison with the other rounds are the key to discovering the mechanics of risk assessment in the context of this and comparable experiments, and perhaps even in real traffic situations.

Another point regarding the present experiment that warrants discussion is the possible anchor effect of the initial speed that each round started with. As indicated in chapter

2.2, the maximum possible speed was 920 pixels per second, which participants did not seem to speed up to. There are two possible explanations for this: Participants might have no longer felt comfortable in their ability to safely dodge the meteors once they flew faster than a certain speed, so they stuck with a speed lower than that. This is reasonable, as at maximum speed, the game was indescribably difficult to play safely even for the very experienced developers. Alternatively, it is possible that the speed each round started with served as an anchor for the speed that participants felt they were "supposed" to play the game at. As every round started with the minimum speed of 320 pixels per second, this anchor is very low. Future experiments should investigate the potential anchoring effect of initial speed by either setting the beginning speed of each round to the maximum speed (920 pixels per second) or by randomizing the speed at which a round starts. If significantly higher average speeds are found in an experiment that lets the ship start with the maximum speed, an anchoring effect of initial speed on subsequent behavior can be inferred. If there is such an anchor effect, the range of speed chosen should increase with the introduction of randomized starting speed. If there is no anchor effect, both findings just described should not appear in subsequent data. Another possibility would be to let the rounds start with a speed of 0 pixels per second and to allow participants to speed up from a dead stop. In order to avoid the data bias described earlier, logging should only start once participants have dodged at least one meteor, as this would require them to have sped up and would have required the game to have generated enough meteors that one of them has already become a danger.

This experiment implies, albeit only through one variable, that risk compensation is real in situations of imminent danger (as implied by the changes found in maximum speed). It was shown that an increased amount of perceived safety causes people to adapt their behavior to be more risky, but only during very short, exceptionally dangerous moments. However, this increase in risky behavior did not cause a higher number of fatalities within the same timeframe, making true homeostatic processes doubtful. It is entirely possible that, as predicted by Wilde (1982a), homeostatic processes take considerable time to emerge, so the very limited timespan that the game was played in might have been too short to detect risk homeostasis. The fact that risk compensation has been shown in this experiment is further evidence for the limited usefulness of safety precautions in all sorts of applications as they might very well be compensated for by humans. Although homeostasis failed to show up in this experiment, the proposal made by Wilde that only a change in the amount of risk people are willing to take can truly affect the rate of incidents might still very well be true. Future research in this regard should try to manipulate the amount of risk people are willing to take before the experiment to see if a difference in, for example, speed emerges. As showing people the consequences of risky behavior has not succeeded in influencing neither participants' willingness to take risk nor their risk behavior in this experiment, another approach is needed. As mentioned already, maybe the way the consequences of behavior were shown in the videos for this experiment was not strong enough – as humans are social and empathic beings, maybe showing a video of a person playing the game, crashing the ship and being visibly upset about it has a stronger effect on willingness to take risk, as people want to avoid feeling like the person they have just been shown. The inverse might be true for showing a person playing carefully and enjoying the game without taking risks. Again, the potential pitfall in this process lies in a possible anchor effect of the behavior shown: instead of influencing the amount of risk people are willing to take, the manipulation just makes a way to play the game more salient, increasing the chance of it being replicated. This could be circumvented by increasing the salience of consequences of behavior in a context that is not comparable to the experiment, for example by showing accidents or safe behavior in traffic or in another domain. This approach could increase or decrease the observer's general desire for safety without making a certain way to play the game more salient. The fact that this type of manipulation needs to be of sufficient strength is evidenced by the fact that the differences in the dependent variables, although all non-significant, are visibly larger between the two video-based groups than between the text-based or control groups.

The present experiment did not test for the influence of psychophysiological arousal in the context of risk-taking. Future experiments should aim to uncover if moments of maximal risk (for example, close dodges or near-misses) cause a response in, for example, skin conductance or hormonal balance to see if higher arousal can be measured in these moments. As discussed in chapter 1.1, this procedure might shed light on whether taking risks is related to a "target level" of psychophysiological arousal and might therefore, for all practical intents and purposes, exist for its own sake rather than for external utility.

Overall, this thesis discussed Risk Homeostasis Theory as introduced by Wilde in 1982, its proposed mechanics with a focus on risk compensation and the theory of gain- and loss-framed messages with their respective potential influence on behavior. The aim of the presented experiment was to uncover possible effects of gain- and loss-framed messages about consequences of actions on risk behavior and to investigate the mechanics of risk compensation as described in Risk Homeostasis Theory (Wilde, 1982a). This was realized in an experiment that featured a controllable spaceship that had to be steered safely, without contact, through a field of approaching meteors. The game traced and logged the speed of the ship, the time it would take the ship to collide with a meteor in its path if it is not steered out of the way and the distance to the closest meteor during gameplay. These parameters were assumed to be indicators of the amount of risk people were willing to take. The game was played over multiple rounds and each round started with a different number of shields that prevented the ship from being destroyed in case of a collision – therefore, a higher number of shields meant a higher amount of objective safety. In accordance with the mechanic of risk compensation proposed in RHT, participants compensated for added safety by behaving more riskily, but only during short moments where their speed was maximized and not during the entire gameplay period. Regardless of the change in risk behavior, the rounds with more shields took longer before the ship was destroyed, indicating that homeostatic processes through *full* compensation did not take place. Before the experiment, participants received different instructions that either explained or showed the consequences of safe behavior (no collisions) or of risky behavior (too high speed, collisions, destruction of the ship). While it was expected that participants would differ in their risk behavior depending on which instruction they had received, such differences did not emerge as significant. Whether no differences were found due to a weakness in the instructions or whether such different instructions generally have no effect on risk behavior should be subjected to future research. Subsequent studies should also experiment with different amounts of shields for the rounds and with different starting speeds for each round.

Risk Homeostasis Theory remains a topic of considerable debate to the present day. Even though this thesis was able to shed some light on the mechanics of risk compensation, which are an integral part of RHT, many of the underlying workings still remain undiscovered and should be a focal point in future research.

5 References

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

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

Christie, R. (2001). The effectiveness of driver training as a road safety measure: An international review of the literature. In *Road Safety Research, Policing and Education Conference, 2001, Melbourne, Victoria, Australia*.

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

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.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychological review*, *62*(4), 243.

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

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.

Hwang, Y., Cho, H., Sands, L., & Jeong, S. H. (2012). Effects of gain-and loss-framed messages on the sun safety behavior of adolescents: The moderating role of risk perceptions. *Journal of health psychology*, *17*(6), 929-940.

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

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. American psychologist, 58(9), 697.

Leuba, C. (1955). Toward some integration of learning theories: The concept of optimal stimulation. *Psychological Reports*, *1*(g), 27-33.

Maslowsky, J., Buvinger, E., Keating, D. P., Steinberg, L., & Cauffman, E. (2011). Cost-benefit analysis mediation of the relationship between sensation seeking and risk behavior among adolescents. *Personality and individual differences*, *51*(7), 802-806.

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

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

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

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

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. (2005). Risk homeostasis theory and traffic education requirements. In *Proceedings of the IV ICTCT Extra Workshop "Measures to assess risk in traffic as reflected by individual test performance, in attitude measurement and by behaviour and interaction"[online]. Campo Grande, Brasil. Available from:* http://www.ictct.org (date last viewed: 10/5/15).

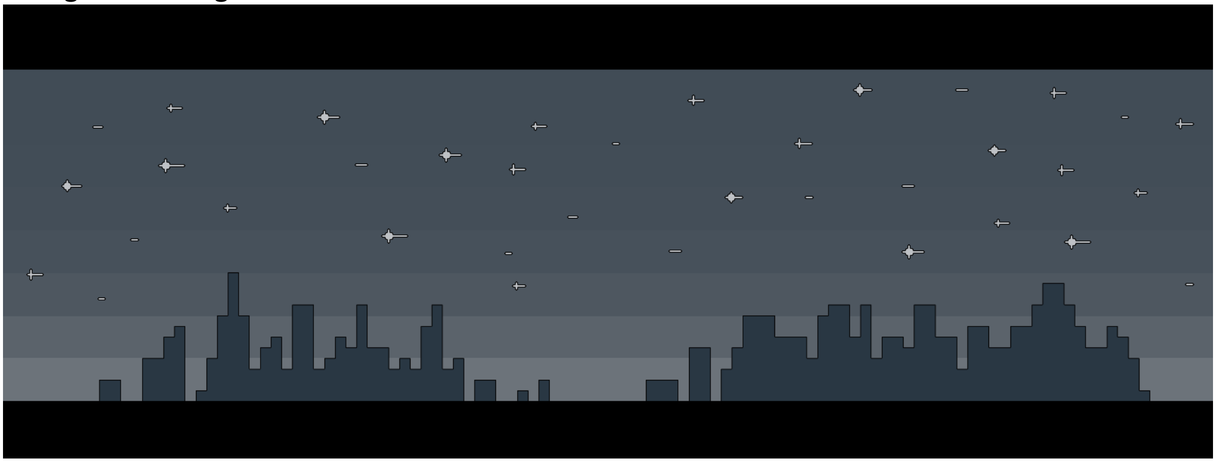
Wilde, G. J., Robertson, L., & Pless, I. B. (2002). Does risk homoeostasis theory have implications for road safety. *BMJ*, *324*(7346), 1149-1152.

Appendix

Spaceship Sprite

Meteor Sprite

Background Image



Example gameplay screenshot