

**Inauguraldissertation
zur Erlangung des akademischen Doktorgrades (Dr. phil.)
im Fach Psychologie
an der Fakultät für Verhaltens- und Empirische
Kulturwissenschaften
der Ruprecht-Karls-Universität Heidelberg**

Titel der Dissertation

Choose a juice!

The effect of choice options, demand and harmful intentions on
aggression in a modified Hot Sauce Paradigm

vorgelegt von
von Dipl.-Psych. Susanne Beier.

Jahr der Einreichung
2012

Dekan: Prof. Dr. Klaus Fiedler
Berater: Prof. Dr. Klaus Fiedler

Acknowledgements

Many people have accompanied the scientific endeavors that culminated in this dissertation. They made these past three years very special and I want to express my deepest thankfulness to them:

I am very thankful to my supervisor Klaus who gave me the opportunity to investigate my own scientific ideas and provided me with critical and enriching feedback for those ideas. I am also thankful to all the research assistants that mixed all kind of juices, were late and rude just as they were supposed to be and, with their help, made my research possible. I also want to thank my crispy colleagues André, Malte, David, Ute, Tobias, Matthias, Max and Momme that made my time at the office a fun time with nerdy, thought-provoking, and enjoyable conversations. Especially, I am very thankful to Florian, who helped me sharpen my scientific ideas and provided me with his inspirational insight in research but also in life, to Philipp, who lent me his ears in countless office buddies conversations that I always look forward to and to Mandy, who not only provided me with an excellent scientific role model but also became a supporting friend. I also want to thank my family for their support.

Summary

One of the main goals of aggression research is to examine the conditions that heighten the likelihood of aggressive behavior. Many studies with a variety of laboratory paradigms have been conducted to address this question. However, many of these studies and the applied paradigms suffer from shortcomings, including a large distance between victim and aggressor, demand characteristics or cues permitting the ostensible aggressive behavior, as well as disregard of participants' affective states, motivations and intentions.

In modern laboratory research some of these shortcomings have been overcome and new paradigms like the *Hot Sauce Paradigm* – a commonly used paradigm in modern laboratory research – have evolved. Paralleling real assault and child abuse cases, in which hot sauce was put in the food of other people or children with the intention to cause them harm, this paradigm measures aggression via the amount of hot sauce participants allocate to another person that allegedly provoked them beforehand. As most paradigms the Hot Sauce Paradigm still has shortcomings that need to be improved. One major shortcoming the Hot Sauce Paradigm shares with some classical paradigms is that participants are not provided with non-aggressive choice options. Without choice options it is impossible to answer questions about conditions under which people choose to act aggressively, since participants do not actually have a choice in regard to their behavior. It is only possible to answer questions regarding the amount of aggression presupposing that participants do act aggressively.

One of the main goals of this dissertation was to analyze the effect of choice options on aggression in a modified Hot Sauce Paradigm. With 5 studies presented in this dissertation, I attempted to answer the overriding question whether or not the

validity of the Hot Sauce Paradigm, as one commonly used paradigm to measure and analyze aggression, could be improved by providing response options (pleasant, neutral, aggressive option) to participants. In general, over the course of five studies I found evidence questioning the traditional paradigm's validity to capture aggressive behavior but also evidence for an enhanced validity of the paradigm when choice options were included.

To test the effect of choice options on the behavior observed in the Hot Sauce Paradigm I used a modified version of the paradigm in the first study. Participants chose one juice out of three juices (pleasant, neutral, spicy juice) to be consumed by a target person that had allegedly just chosen either a sour (provocation) or neutral (no provocation) juice for them. This condition with choice options was contrasted with three control conditions in which participants could only administer different amounts of either the pleasant, neutral or spicy juice. Provoked participants that were provided with a spicy juice only administered more spicy juice than non-provoked participants, replicating previous findings with the Hot Sauce Paradigm. However, there also was a main effect of provocation on the allocated juice with provoked participants generally allocating larger amounts of any type of juice than non-provoked ones. With choice options, critically, only three of the 18 provoked participants did not choose the pleasant juice and chose the neutral (1 participant) or spicy juice (2 participants) instead. None of the 20 unprovoked participants chose the spicy juice. These findings question the original Hot Sauce Paradigm's validity to capture aggressive behavior with an intention to harm another person. The low choice rates for the aggressive option also highlight the importance of providing non-aggressive response options to participants to avoid an overestimation of aggressive and an underestimation of non-aggressive behavior.

I hypothesized that the main effect of provocation on the allocated amount that I found in Study 1 was due to provoked participants' higher responsiveness to the paradigm's demand in the direction of applying more juice resulting from a lowered self-control after experiencing a provoking event. The aim of Study 2 and 3 was to test this assumption. In Study 2, I replicated the result of provoked participants allocating more spicy juice than non-provoked ones in conditions without alternative choice options. However, this effect disappeared with less guiding instructions and thus, again, questioning the traditional paradigms validity. To test whether or not this responsiveness to the demand of the paradigm is associated with the participants' self-control, I measured participants' state self-control in both Study 2 and 3, in which I applied the same provocation method as Study 1, and in Study 4, in which I used a more intense interpersonal provocation. In none of these studies I found a difference in self-control between provoked and non-provoked participants after the provocation. This makes the proposed self-control explanation for the allocation amount results unlikely.

The amount of aggressive choices shown in Study 1 was relatively low (2 of 18 provoked participants chose the spicy juice). To test whether a more intense form of provocation would increase aggressive choices, which would make additional analyses and ultimately the application of the modified paradigm in aggression research possible, I used a more intense interpersonal provocation in Study 4. The provocation included waiting for another alleged participant and rude behavior toward the actual participant shown by the unpunctual confederate. To also test the effect of victim visibility – another shortcoming of modern laboratory research – on the behavior shown in the Hot Sauce Paradigm, I orthogonally manipulated the visibility of the victim with the provocation manipulation. To assure that the measured behavior

captures aggression, I measured the key aggression defining variable harmful intentions of participants. The results of Study 4 indicated that descriptively aggressive choices occurred more often than in Study 1 and that provoked participants chose the spicy juice significantly more often than non-provoked ones. This effect of provocation on spicy juice choices was mediated by the participants' harmful intentions indicating the validity of juice choices as a measurement of behavioral aggression. The visibility of the victim was not significantly associated with differences in the juice choices.

To gain further evidence for a connection of juice choices and harmful intentions and to externally validate juice choices as aggression measures I analyzed the perception of intentions behind juice choices from the victim's point of view in Study 5. In a scenario study, modeled after Study 4, participants imagined to receive a sample of the spicy juice from another participant in a study for which they arrived 15 minutes late. Intentions behind the juice allocation were perceived as more harmful if the person who made the allocation was provided with choice options in comparison to if the person was not provided with choice options (like in the traditional Hot Sauce Paradigm). This finding supports the claim that the inclusion of choice options heightens the paradigm's validity, since the connection with a key aggression defining construct was strengthened with choice options in comparison to without them.

The present research underlines the importance of providing non-aggressive choice options to participants in laboratory aggression research. Particularly if the conditions under which aggressive behavior occurs are examined, participants have to be able to make a choice of what kind of behavior they want to show. With the modified Hot Sauce Paradigm introduced in this dissertation a research instrument is provided to achieve this goal.

Content

1	Introduction.....	8
	1.1 Laboratory aggression research.....	10
	1.2 Choice options.....	18
	1.3 Additional shortcomings.....	26
2	Overview.....	28
3	Study 1.....	28
	3.1 Method.....	29
	3.2 Results.....	34
	3.3 Discussion.....	39
4	Study 2.....	42
	4.1 Method.....	43
	4.2 Results.....	44
	4.3 Discussion.....	47
5	Study 3.....	49
	5.1 Method.....	49
	5.2 Results.....	51
	5.3 Discussion.....	53
6	Study 4.....	54
	6.1 Method.....	56
	6.2 Results.....	60
	6.3 Discussion.....	66
7	Study 5.....	67
	7.1 Method.....	68
	7.2 Results.....	70
	7.3 Discussion.....	70
8	General Discussion.....	71
	8.1 Summary of the results.....	71
	8.2 Limitations and further research proposals.....	79
9	Conclusion.....	81
	References.....	81
	Appendix 1.....	92
	Appendix 2.....	94
	Erklärung gemäß § 8 Abs. 1 Buchst. b) und c) der Promotionsordnung.....	
	der Fakultät für Verhaltens- und Empirische Kulturwissen.....	97

Introduction

“But I tell you, Do not resist an evil person. If someone strikes you on the right cheek, turn to him the other also” (Matthew 5:39)

These pacifist words of Jesus illustrate that humans are not bound to retaliate or seek revenge via aggressive behavior. Nevertheless, under certain conditions some humans show aggressive behavior. Most researchers would agree on the following definition according to which aggression is defined as an intentional attempt to harm another person (Dollard, Doob, Miller, Mower, & Sears, 1939; Berkowitz, 1962). More precisely, Baron and Richardson (1994, p. 7) describe aggression as “any form of behavior directed toward the goal of harming or injuring another living being who is motivated to avoid such treatment.”

Aggressive behavior occurs in a multitude of ways (Krahé, 2001). For example, aggression can be verbal or physical behavior. It can occur as a direct attack, or in a more indirect way without direct contact, and thus aggression can occur either in a more overt and obvious manner or in a covert more disguised form (e.g., Berkowitz, 1993; Krahé, 2001; Ritter & Eslea, 2005). Also, aggressive behavior can occur impulsively or in a more consciously controlled way (e.g., Berkowitz, 1993). Aggressive behavior can be shown with the main objective to harm someone (hostile aggression) and without a provocation (proactive aggression) or with additional objectives besides causing harm to another person (instrumental aggression) as retaliation after an initial provocation (reactive aggression; e.g., Berkowitz, 1993; Geen, 2001; Krahé, 2001). Aggression can be carried out by individuals or by groups.

Over decades of research thousands of experiments in psychology using a multitude of different paradigms have been performed trying to answer one core

questions of aggression psychology: Under which conditions do humans show aggressive behavior? The variety of methods to study aggression can be divided in two general approaches: observing and asking (Baron & Richardson, 1994). Aggression can be observed in naturalistic contexts in which the frequency of specific forms of aggression can be assessed but it can also be observed in field and laboratory experiments (Krahé, 2001). Besides the advantage of assessing aggression in the social context in which the behavior occurs with all the complexity and diversity that characterizes real-world aggression (Graham, Tremblay, Wells, Parnanen, Purcell, & Jelly, 2006), lots of variance is not under the experimenter's control in field experiments. Laboratory experiments, in contrast, allow drawing causal inferences (Krahé, 2001). Observational methods are not always possible or feasible. A researcher can instead base his or her analysis on archival data (e.g., crime statistics) that were collected for other reasons (Krahé, 2001). Another method for studying aggression is to obtain self-reports about aggressive behavior or about stable differences in cognitions and affects associated with aggressive behavior, for example, a self-report of physical aggression or anger measured with Bush and Perry's (1992) "Aggression Questionnaire" (e.g., Krahé, 2001). Responses on these measures, though, are influenced by people's tendency to appear in a socially desirable way. Implicit measures of aggression try to overcome this bias by relying on automatic cognitive and affective responses. Implicit aggression is, for example, assessed via speeded reactions in a double categorization task with categories like aggressive versus peaceful behavior and pleasant versus unpleasant (Bluemke & Zumbach, 2007). Also less susceptible to biases due to social desirability are techniques that assess a person's aggressive behavior or general aggressiveness via nomination by peers or by others or with other-reports. However, these techniques can

be very time-consuming and the convergence over persons is necessary to reliably assess the behavior of the target person (Krahé, 2001).

In this dissertation I focused on measurement of aggression in the laboratory. The major research goal of my dissertation was to improve the validity of a commonly used laboratory research paradigm, the so-called Hot Sauce Paradigm (McGregor, Lieberman, Greenberg, Solomon, Arndt, & Simon, 1998; Lieberman, Solomon, Greenberg, & McGregor, 1999), by providing participants with non-aggressive response options missing in the original paradigm. In my dissertation I will demonstrate theoretically as well as empirically that in laboratory research aggression is not only operationalized more appropriately for an analysis of conditions of aggressive behavior when choice alternatives are given to the participants. The inclusion of choice alternatives in the Hot Sauce Paradigm also enhances the link between observed aggressive behavior and harmful intentions which is one of the key aggression defining theoretical concepts. Thus, the validity of the paradigm is improved by an inclusion of choice alternatives.

1.1 Laboratory aggression research

Over decades of laboratory aggression research, provocation has been identified as an important trigger of any form of aggressive behavior (e.g., Anderson & Bushman, 1997; Carlson & Miller, 1988). Other examples of situational variables demonstrated to heighten aggression are alcohol (Anderson & Bushman, 1997; Ito, Miller, & Pollock, 1996) and the presence of stimuli cuing aggressive behavior (e.g., a picture of a gun; Berkowitz & LePage, 1967; Anderson, Benjamin, & Bartholow, 1998) especially when these variables are combined with a form of provocation or frustration (Krahé, 2001). Laboratory research not only identified aggression-eliciting factors like the ones stated above but also investigated moderators for these

relationships. The link between aggression cues and aggressive behavior is moderated, for example, by the person's sex. Irrespective of provocation, a connection between aggression cues and aggressive behavior was demonstrated for men. For women, however, aggression cues heighten the aggression level only in the presence of a provocation (Krahé, 2001).

Many of these experiments that addressed the core issue of conditions under which aggressive behavior is facilitated have used one of the following classical laboratory aggression paradigms: 1) Teacher-Learner Paradigm, 2) Essay Evaluation Paradigm, 3) Competitive Reaction Time Task, and less often 4) Bobo Doll Modeling Procedure (Tedeschi & Quigley, 1996). The naïve participant in the Teacher-Learner Paradigm (Buss, 1961) always teaches a confederate in a memory task by punishing incorrect responses that the confederate gives at predetermined times usually by delivering electric shocks. In a newer version of the paradigm, other aversive stimuli, e.g., noise blasts are used instead of shocks (Krahé, 2001). Aggression is measured via the frequency and intensity of the punishment in this paradigm (Tedeschi & Quigley, 1996). The general logic behind the Essay Evaluation Paradigm (Berkowitz, Corwin, & Heronimus, 1962) is similar. However, instead of taking the role of a teacher, participants in this paradigm are asked to evaluate an essay allegedly written by another participant via the application of electric shocks. Aggression, again, is measured via the number of electric shocks (Krahé, 2001). Both these paradigms provide participants with cover stories, making them believe that they either teach or evaluate another participant. As a consequence, they might not apply the aversive stimulus with the goal to harm the other person but with a prosocial goal that is in line with the provided cover stories (Tedeschi & Quigley, 1996). Referring to the definition of aggressive behavior as actions that are carried out with the intention to

harm someone (Baron & Richardson, 1994), it is highly questionable whether both paradigms provide valid measures of aggression. In addition, these paradigms suffer from other shortcomings like an experimenter who serves as an aggression approving authority figure as well as a large spatial distance between the confederate and the participant since they are located in different rooms which leads to artificial situations with potentially different results if the distance between victim and aggressor was lower (Tedeschi & Quigley, 1996). As Ritter and Eslea (2005) concluded, it is therefore most likely that the behavior shown by participants in laboratory experiments conducted with these paradigms is very different from real-life aggressive behavior.

In response to some of the critical issues of the Teacher-Learner Paradigm and the Essay Evaluation Paradigm especially the potential prosocial motive for the application of aversive stimuli the Competitive Reaction Time Task (also called Taylor Aggression Paradigm; Taylor, 1967) was developed (Giancola & Chermack, 1998). In this task participants engage in a competitive reaction time game with another alleged participant. Aggression is measured via the shock or noise blast intensity delivered to the opponent on winning trials. More precisely, participants are instructed to show a specific reaction, for example, pushing a button as fast as possible. Before each trial participants set the intensity of the shock or noise that the slower one of the two will receive after the trial. The application of shocks/noise blasts to the participant and the confederate after the trials actually follows a schedule predetermined by the experimenter. One of the major advantages of this paradigm is that no prosocial cover story is used (Tedeschi & Quigley, 1996). However, shortcomings like a large unvaried distance between the aggressor and the victim as well as an aggression approving experimenter still remain unresolved. In addition,

participants might intensify shocks due to the competitive nature of the paradigm and because they are motivated to win the competitive game (Tedeschi & Quigley, 1996, 2000; Ritter & Elsea, 2005).

The last paradigm occasionally used in classical laboratory research especially with children is the Bobo Doll Modeling Procedure (Bandura, 1973). In this paradigm participants first observe a model engaging in an aggressive play with a large inflatable doll. Then they themselves have the opportunity to play with that doll. Here, the primary measure for aggression is the amount of aggressive behavior shown by participants. However, participants' behavior might be no more than joyfully playing in a rough and tumble way without any intention to harm someone (Tedeschi & Quigley, 1996).

Next to these paradigms focusing on physical aggression, verbal aggression has been assessed in the laboratory. In the typical study participants verbal evaluation are assessed and aggression is measured via the aggressive content in these evaluations. In order to qualify as an aggression measure, participants have to believe that their negative evaluations lead to harmful consequences for the evaluated person (Baron Richardson, 1994; Krahe, 2001).

Despite these problematic issues of classical laboratory research paradigms, previous research that attempted to answer the question if laboratory aggression research, often conducted with one of the discussed paradigms, assesses a common construct, demonstrated that different measures of physical aggression (e.g., shock intensity, shock duration) as well as effect-size estimates of written and physical aggression were positively correlated. Also, verbal and physical aggression were both influenced similarly by aggression eliciting factors, discussed in the literature (frustration, anger, directness of aggression, exception: physical attack; Carlson,

Marcus-Newhall, & Miller, 1989). Thus, these authors interpreted these results as evidence for the construct validity of the paradigms used in laboratory aggression research and that all these paradigms capture the common construct aggression. However, as Tedeschi and Quigly (1996, 2000) pointed out, in spite of these positive correlations and similar eliciting factors pointing toward the measurement of a common construct, this measured underlying construct might be somewhat different from aggression, for example, compliance to cues indicating participants how the experimenter expects them to behave, conformity (Gottfredson & Hirschi, 1993) or a negative reciprocity in the sense that participants just match the level of the noxious stimulus from their opponent (Tedeschi & Quigly, 1996, 2000).

In a similar vein Anderson and Bushman (1997) argue in defense of the validity of laboratory aggression research mostly conducted with either the Teacher-Lerner Paradigm, the Essay Evaluation Paradigm or the Competitive Reaction Time Task that, in spite of the critical aspects raised by Tedeschi and Felson (1996), real-world and laboratory studies show similar results thus supporting the validity of laboratory studies as well as the used paradigms. More precisely, they argue that both studies conducted inside and outside the laboratory indicate that men show more aggression than women, that trait aggression is associated with an increase in aggression, that provocation increases aggression as well as alcohol consumption and anonymity. However, this convergent evidence might still be ambiguous in regard to the validity of laboratory aggression research paradigms. Concerning anonymity, for example, conflicting research exists indicating that anonymity can lead to both an increase in aggression as well as prosocial behavior depending on salient cues available in a situation (for a meta-analytic review see Postmes & Spears, 1998). Johnson and Downing (1979) ran a Teacher-Lerner Paradigm study in which they

manipulated the anonymity of the participants. They also varied the way in which participants were deindividuated. In one condition participants were made non-identifiable similar to Zimbardo's (1970) study by making them wear lab coats and huts resembling Ku-Klux-Klan dresses. Zimbardo (1970) showed an increase in aggression under such an anonymity manipulation in his study that he interpreted as an increase in aggression due to a state of deindividuation. Johnson and Dowing (1979), however, also included a condition in which participants were deindividuated by making them wear nurse uniforms. These authors showed that participants in the Ku-Klux-Klan like dresses condition increased shocks in the course of the study and participants in the nurse uniform condition decreased the shock level. Similar to this exemplary study, Hirsh, Galinsky, and Zhong (2011) reviewed the literature on behavioral consequences of alcohol consumption, power and anonymity and demonstrated with reference to their proposed General Model of Disinhibition that the consequences of these variables can either be pro- or anti-social depending on salient cues available in a specific situation. Thus, the convergent evidence reviewed by Bushman and Anderson (1997) that I discussed above can also be interpreted as being caused by a specific research focus on aggression in comparison to, for example, prosocial behavior in both real-world aggression studies and demand cues in laboratory studies increasing the salience of aggressive behavioral responses instead of indicating the validity of the paradigms used to study aggression.

To overcome some of the aforementioned critiques of classical laboratory research and differences between laboratory aggression measures and real-life aggression in regard to surface characteristics (Anderson, Lindsay, & Bushman, 1999), improved paradigms have been proposed more recently. In the Point Subtraction Aggression Paradigm (Cherek, 1981), participants earn money by either

pressing a specific button or by subtracting it from another alleged participant (aggression measure). While in this paradigm participants can also show non-aggressive behavior, aggression is the only way of interacting with the other person (Tedeschi & Quigly, 2000). In the Bungled Procedure (Russell, Ams, Loof, & Dwyer, 1996), aggression is measured via the power of a gun and the number of pellets selected by participants to ostensibly shoot at a female target in an alleged new game. The external validity in this paradigm was improved since the task resembles real-life aggression more closely. However, participants never actually engage in the aggressive behavior in this paradigm (Ritter & Eslea, 2005). Another attempt to improve the external validity is the Experimental Graffiti and Tearing Procedure (Norlander, Nordmarker, & Archer, 1998). In this paradigm pictures are given to participants and they are asked to destroy them by drawing on them and tearing them apart. Aggression is measured via the rated amount of destruction of the drawings and the number of pieces resulting from tearing them apart. Thus, participants are specifically instructed to show the aggressive behavior and the paradigm permits aggression (Ritter & Eslea, 2005). Another paradigm working with pictures is the IAPS-Picture Selection Task (Mussweiler & Förster, 2000). Aggression is measured via the average valence of pictures allegedly selected for another participant to look at in a study from a collection of positive, neutral and negative pictures. However, in comparison to the previous paradigm, the resemblance to real-life aggression is less obvious. Another paradigm that has been proposed more recently is the Hot Sauce Paradigm (McGregor et al., 1998; Lieberman et al., 1999). This paradigm in particular has many advantages compared to classical paradigms and is commonly used in modern laboratory aggression research covering a large variety of topics, including, for example, the link between self-regulation and aggression (DeWall, Baumeister, Stillman, & Gailliot, 2007), violent video games and aggression (e.g., Barlett, Branch,

Rodenheffer, & Harris, 2009; Adachi, & Willoughby, 2011) or ostracism and aggression (Warburton, Williams, & Cairns, 2006). Participants in studies conducted with this paradigm are instructed to determine how much hot sauce another alleged participant with a distaste for spicy food has to consume. Aggression is measured via the amount of the administered hot sauce. More precisely, participants are told that they participate in a study that analyses the relationship between personality and taste preferences. They are instructed to prepare a hot sauce sample for another alleged participant in an adjacent room that had either just provoked them in some form or not. Before preparing the sample participants receive a questionnaire allegedly indicating the target's general taste preferences including the information that this person does not like spicy food.

In comparison to the classical experimental paradigms one main advantage of the Hot Sauce Paradigm is the improved external validity since real-world assaults involving the use of spicy food to harm someone have been documented in the past. Hot Sauce, for example, has been fed to children by caretakers in child abuse cases (e.g., Regan, 2012; Huffington Post, 2012; Stogsdill, 2012) and in assaults, for example, by fraternity students forcing others to drink Tabasco Sauce (Lee, 2009). Additional advantages of the paradigm are that expensive materials like an aggression machine allowing the application of electric shocks are not necessary (cf., the Teacher-Learner Paradigm is conducted with a Buss aggression machine necessary for the application of electric shocks), the dependent variable is quantifiable as well as easy to measure and the observed behavior is neither biased due to competitive motives nor due to a prosocial cover story (Ritter & Eslea, 2005). Thus, the Hot Sauce Paradigm can be regarded as a considerable improvement over classical laboratory paradigms.

1.2 Choice options

The Hot Sauce Paradigm, as most other paradigms, is not perfect and can still be improved further. As criticized for the classical paradigms by Tedeschi and Quigley (1996), the classical aggression paradigms reviewed above (excepting the Bobo Doll Modeling Procedure), which allow for alternative explanations other than aggression, often suffer from an additional shortcoming that they have in common with the Hot Sauce Paradigm: They often do not provide any response alternatives other than an aggressive one to participants. Thus, in these cases participants can only vary the amount of supposedly aggressive behavior. However, when paradigms do not include non-aggressive choice options it is not possible for participants to react in a non-aggressive way (Tedeschi & Quigley, 1996; Ritter & Eslea, 2005).

For the classical paradigms like the Competitive Reaction Time Task some improvements have been made. Some studies with the white noise version of the Competitive Reaction Time Task include the option to apply no noise (e.g., Bushman, 1995). To set the noise level, participants thus are provided with a scale that, for example, ranges from 0 to 10 as options for the noise level.

Even if such an explicit no shock/no noise option does not exist, it sometimes has been argued for the classical paradigms that the lowest possible option, which can be selected for the noxious stimulus, can be regarded as a non-aggressive response since the participants are informed that the lowest option results only in a mild noxious stimulation (Giancola & Chermack, 1998). In line with this argument, participants indicate that they selected the lowest setting for the noxious stimulus in order to not hurt the alleged other participant (Taylor & Leonard, 1983). However, even though some participants try to find a way in these paradigms to cause as less harm as possible for the alleged other participant, it does not mean that this is an ideal

solution that does not need to be improved further. Especially if there is no zero option available, the aggressive option is not qualitatively different from a low noxious stimulus application, but only quantitatively. Providing qualitatively different response options that every participant can recognize as such might affect the participants' behavior and might lead to different results in comparison to when such options are not provided. For example, in their meta-analysis on the effect of alcohol on aggression Ito et al. (1996) found stronger effects of alcohol consumption when participants were not provided with choice options in comparison to when they were provided with choice options. The lack of (qualitatively different) choice options might also affect the paradigms' validity due to demand characteristics. If participants are only provided with one response alternatives and they can only vary the quantity of the behavior, it is conveyed to participants that the experimenter allows this behavior, e.g., the application of noise, shocks etc.. Such an experimental setting might signal participants that they are expected to show this behavior. Also, the availability of only one option functions as a permissive cue for the aggressive behavior (Ritter & Eslea, 2005).

In newer paradigms like the Point Subtraction Paradigm participants have other options besides aggressive behavior. In this paradigm participants allegedly play a computer game with an opponent. Participants have two response options. They can either receive a point by pressing button 1 multiple times (e.g., 100 times) or they can steal a point from their opponent by pressing button two multiple but not as many times as button 1 (e.g., 10 times; e.g., Cherek, 1981; Giancola & Chermack, 1998). In newer versions a third button sometimes exists. Pressing this button protects the participants' points from the opponent for a certain time interval in which there are no provocations (Geniole, Carré, & McCormick, 2011). Aggression is measured via the

number of times points are stolen from the opponent. However, even in this paradigm, especially in the older versions of it, the only way of interacting with the other participant is by stealing points since this is the only way of engaging with the alleged opponent and possibly influencing his/her behavior (Tedeschi & Quigley, 2000). In the newer version, it might be that participants chose to protect their points also to communicate a peaceful strategy to the alleged opponent. However, they also might choose this option, for example, because they are anxious or they want to withdraw from the interaction.

Due to improvements in regard to the external validity and no potential bias caused by competitive or prosocial motives (Ritter & Eslea, 2005) the Hot Sauce Paradigm can be regarded as an improved measure in comparison to classical paradigm. However, in the Hot Sauce Paradigm aggression is measured via the allocated amount of hot sauce. Thus, participants can only vary the degree of aggression. They are not provided with choice alternatives besides the aggressive behavior if allocating no sauce at all is not considered a real choice option (Ritter & Eslea, 2005) due to the above mentioned reasons. If these alternative behavioral options are not provided the only questions that can be answered concern the intensity of someone's aggressive behavior under certain conditions given that the person does react with aggression. The more interesting research questions of whether or not someone reacts aggressively or which conditions heighten the likelihood of the occurrence of aggressive behavior cannot properly be answered without providing choice alternatives. In real-life situations showing aggressive behavior is only one of many options humans can choose from (Tedeschi & Feslon, 1994). Instead of behaving aggressively, people might, for example, leave a provoking situation instead of retaliating or might even try to calm the aggressor down by doing something nice

for him/her and the like. Without providing alternative behavioral options to participants, conclusions about how they would have reacted with these options are impossible (Ritter & Eslea, 2005).

If researchers provide only an aggressive response option they expect that a person would choose aggressive behavior in a situation to which the results of the study would apply. A critical test of whether or not certain conditions that might be associated with aggression also enhance the likelihood of other types of behavior (see, for example, Hirsh et al., 2011) is not conducted if participants are not provided with non-aggressive response options. Focusing the information search for testing a hypothesis only on the behavior that is expected to occur while other forms of potential outcomes are disregarded inherits a confirmation tendency (Fiedler, 2011) and can be classified as positive hypothesis testing strategy (Klayman & Ha, 1987; Fiedler, Walther, & Nickel, 1999). Other forms of behavior elicited in situations typically associated with heightened aggression that might well be prosocial or pacifist are systematically underestimated with such a research strategy. The above cited classical example of Jesus suggesting to offer the other cheek after being slapped in the face is only one of many examples in which an aggressive reaction might be the expected behavior but still (some) people would not choose to act aggressively if they have other options. Instead of acting aggressively by putting hot sauce in the food for children parents that are provoked by their children's behavior, for example, might offer positive incentive for stopping the behavior or behave neutrally and ignore the child. As discussed in the previous section, Hirsh et al. (2011) recently proposed and demonstrated that variables that were elsewhere discussed as variables associated with the occurrence of aggressive behavior (e.g., anonymity, alcohol; Anderson & Bushman, 1997) actually foster both aggressive and prosocial

behavior. To detect such behavioral consequences, other behavioral option next to aggressive ones should be provided. With response options, like they exist in real life, conclusions about when and under which conditions people act aggressively would be possible.

Another example in the field of aggression research that suffered from a comparable confirmation tendency and recently started to overcome this flaw is the research on the effects of video games (Ferguson, 2007). Possibly due to the researchers' assumption of negative effects resulting from playing video games that might be due to events like school shootings, in which people want to blame the unexplainable violence on some cause, the vast majority of research in this field focused on demonstrating negative effects, e.g., heighten aggressive feelings and cognitions, aggressive behavior and desensitization to violence due to playing those games (Ferguson, 2007; Barlett, Anderson, & Swing, 2009; Greitemeyer & Oswald, 2012). Possible positive outcomes of violent or prosocial games have largely been ignored in past research (Greitemeyer & Oswald, 2012) and thus unbiased conclusions about the effects of video games were not possible. This confirmation tendency is even heightened by a publication bias in favor of studies that successfully demonstrated a link between negative outcomes and exposure to violent video games (Ferguson, 2007). Recently, however, there has been a bit of an improvement in this field. Positive effects of (violent) video games like an improvement of perceptual accuracy and visual attention have been demonstrated (Barlett et al., 2009). In addition, positive effects of exposure to prosocial games like an increase in helping behavior has been shown (Greitemeyer & Oswald, 2012). Other possible positive effects of (violent) video games might, for example, be an enhancement in subjective well-being or, for certain games, an improvement in leadership and team cooperation

skills (Barlett et al., 2009). Just like research on the effects of video games started to overcome its confirmation tendency, new paradigms of laboratory aggression research like the Hot Sauce Paradigm should also account for choice options existing in real-life by providing choice alternatives. This then would allow unbiased conclusions about the conditions of aggressive behavior.

Tentative evidence from previous research with the discussed classical paradigms indicated changes in participants' behavior when non-aggressive choice alternatives are provided to them and demonstrated that aggressive behavior might not necessarily be the dominant choice of the participants. For example, in a study in which the Teacher-Lerner Paradigm (Buss, 1961) was used to examine the effect of non-insulting aversive events on aggressive behavior Berkowitz, Chochran, and Embree (1981) included the options of rewarding a target with money next to the option of punishing it with noise blast. The choice rates of rewards were much higher in comparison to punishments over all conditions. Similar results were found by Hokanson (1970) in a study in which female participants provided with the response options of administering either a punishing electric shock or a rewarding token to a target that just provoked them chose the friendly option more often than the aggressive one. Also, as mentioned above Ito et al. (1996) found in their meta-analysis on the effect of alcohol consumption on aggressive behavior that the effect was smaller with choice options provided to participants in comparisons to without choice option. In this meta-analysis these authors' analyzed 49 experimental studies that mostly were conducted with the discussed classical paradigms especially the Teacher-Lerner Paradigm and the Competitive Reaction Time Task.

For the Hot Sauce Paradigm some attempts have been made in previous research to overcome the shortcoming of missing choice options. To analyze short

term effects of violent video games on aggressive behavior Barlett et al. (2009), for example, modified the Hot Sauce Paradigm through the inclusion of response options. Instead of receiving only one hot sauce, participants were provided with four hot sauces to choose from with varying degree of spiciness. After tasting all four of the sauces participants were asked to choose one of them and then apply an amount of their choice of the selected sauce in a cup for the alleged other participant. However, even though participants were provided with 4 response options and the provided sauces varied in their degree of spiciness, all four sauces were spicy sauces. Thus, all response alternatives were aggressive ones, yet to a varying degree. Even with the provided choice options non-aggressive response alternatives, e.g., a neutral or pleasant choice alternative, were again missing in this study.

Instead of using only hot sauce, Böhm and Streicher (2010) provided their participants with different kinds of food stimuli. Participants were instructed to allocate an amount of their choice of hot sauce, mayonnaise (indirect form of aggression) or low fat yoghurt (neutral option) in a cup for a target. They were free to choose which one of the food samples to allocate, allocate more than one sample or not allocate food at all. The selection of food stimuli was chosen by these authors because their main research focus was to measure sex differences in direct and indirect forms of aggressive behavior. In line with their hypotheses, Böhm and Streicher (2010) demonstrated that women used indirect forms of aggression more often than men, and thus, with provocation, applied more mayonnaise than without provocation. Through providing more food items than just one hot sauce, the participants' attention focus was no longer only on the aggressive option, and possible effects of the provocation on other types of behavior were detectable as well. At the same time though, these authors did not include a pleasant alternative for the

participants to choose from due to their research focus. Hence, participants were only able to prevent harm by choosing the neutral alternative or by not choosing any alternative at all. They were not able to assign something positive to the target. The choice rates for the provided food samples might well be affected by an inclusion of such an alternative.

These attempts to improve the Hot Sauce Paradigm by adding choice options are very positive. However, it might be possible to further improve the paradigm by providing participants with an aggressive, a pleasant and a neutral response option. One goal of my dissertation research was to analyze the effect of choice options on aggression measured with the Hot Sauce Paradigm as one of the most commonly used paradigms in modern laboratory aggression research. With this goal in mind, I provided participants of the present studies with two behavioral options in addition to the aggressive one – a pretested neutral and pleasant option.

As mentioned in a previous section, besides leading to a confirmation tendency, the lack of choice options weakens the validity of the paradigm due to demand characteristics. If participants are only provided with one response option it signals them that the experimenter allows this behavior, e.g., in the Hot Sauce Paradigm the allocation of the spicy food to the target. This functions as a permissive cue for the aggressive behavior (Ritter & Eslea, 2005). In addition, the demand is further increased in the Hot Sauce Paradigm by instructing participants “to put as much or as little” (Lieberman et al., 1996, p. 339) sauce in the cup for the alleged other participant, which specifically focuses on the amount. This might guide the participants’ behavior toward someone especially when that person just provoked them in an unfamiliar situation. It also cues using the allocated amount as a way of dealing with the situation instead of relying on a spontaneously chosen strategy that

the participants come up with themselves (Ritter & Eslea, 2005). Furthermore, it signals participants a permission to apply a large amount of the juice. With choice alternatives for participants both the confirmation tendency as well as the demand would be lowered since participants' attention focus would no longer be drawn to the one aggressive option. In the research of my dissertation, I contrasted conditions with and without choice options (Study 1, Study 5) and analyzed the effect of guiding cues in the Hot Sauce Paradigm instructions (Study 2).

1.3 Additional shortcomings

Another shortcoming of laboratory aggression research that has not yet been overcome with the Hot Sauce Paradigm refers to the unvaried large distance between the aggressor and the alleged victim. Usually participants in the Hot Sauce Paradigm are told that the target, for whom they prepare the taste sample, is sitting in an adjacent room. Hence, the aggressor does not see the victim in the classical version of the Hot Sauce Paradigm. Taking Milgram's (1974) obedience research and replications with the Teacher-Lerner Paradigm (Ahmed, 1979; Page & Moss, 1976) into account, which demonstrated lower obedience/aggression levels when participants saw the victim, victim visibility might also influence aggressive behavior within the Hot Sauce Paradigm. Thus, it is worth analyzing how the victim's visibility affects participants' behavior in the Hot Sauce Paradigm. In the research of my dissertation I attempted to analyze the effect of victim visibility on aggressive behavior. In Study 4 I manipulated the visibility of the target orthogonally to an interpersonal provocation in order to shed some light on the question of whether or not aggressive behavior in the Hot Sauce Paradigm is influenced by this factor.

Besides these structural shortcomings, there are additional shortcomings in the application of the Hot Sauce Paradigm. Previous research has not sufficiently

analyzed mediating variables like affect, motivation and intention in the Hot Sauce Paradigm (Ritter & Eslea, 2005). In some studies with the Hot Sauce Paradigm the affective state of participants was measured. McGregor et al. (1998), for example, demonstrated that participants' hostility level was higher after they drank a sour (provocation) in comparison to neutral drink (no provocation) allegedly given to them by another participant. In another study focusing on anger a significant increase in anger as well as sadness after negative feedback (provocation) in comparison to positive feedback was demonstrated (Evers, Fischer, Mosquera, & Manstead, 2005). Thus, both studies indicated a successful provocation reflected in changes in participants' affective state. However, since the affective state was measured only once it is unclear how the affective state of participants developed during the course of the paradigm (Ritter & Eslea, 2005). For example, it is unclear if the affective state changed again after participants were given the chance to revenge back by allocating the spicy sauce to the other participant. Motivation and intention are other potential mediators that have not been analyzed thoroughly in the Hot Sauce Paradigm (Ritter & Eslea, 2005). Disregarding these variables is especially problematic since the intention to harm another person is a key element in the definition of reactive aggression (Baron & Richardson, 1994). In their review of current laboratory research Ritter and Elsea (2005) call for a more detailed analyses of the motives behind the sauce allocation that go beyond an analyses of vague statements given by participants during the debriefing sessions about whether or not they wanted to harm the target with the juice allocation. Thus, in the research of my dissertation I measured participants' affective state (Study 1) and their intentions behind the juice allocation (Study 2, 3, 4) as well as the perceived intentions of a victim in the classical and the modified Hot Sauce Paradigm that included choice options (Study 5).

2 Overview

In this dissertation I analyzed effects of major shortcomings of classical and current laboratory aggression research using the Hot Sauce Paradigm as one of the commonly used modern laboratory aggression research paradigms. In a series of 5 studies I examined the effects of including additional non-aggressive choice options (Study 1, 4), manipulating demand characteristics (Study 1, 2), and varying victim visibility (Study 4) on aggression in the Hot Sauce Paradigm. To ensure that the behavior shown by participants qualifies as aggression I measured participants' intentions to harm (Study 2, 3, 4) and analyzed the perception of the allocators' behavior by victims receiving the juice sample (Study 5). In sum, I found evidence questioning the validity of the classical Hot Sauce Paradigm as well as evidence for a considerable improvement of the validity of the modified paradigm due to the inclusion of choice options.

3 Study 1

With the first study I aimed to analyze the effect of multiple choice options in a modified Hot Sauce Paradigm (McGregor et al., 1998). Critically, I extended the Hot Sauce Paradigm by including two additional response alternatives – a pleasant and a neutral choice option – next to the commonly used aggressive one. Participants chose one option to allocate to a target person from whom they allegedly had just received a sour (provocation) or neutral (no provocation) juice themselves. In addition to this condition with choice options, I included three control conditions with no choice options. Depending on the experimental condition, participants in the groups without choice options could only decide about the amount of either the pleasant, neutral or aggressive option. Thus, with all these conditions a comparison of participants' behavior in conditions with and without choice options was possible. If the amount of hot sauce allocated in the classical Hot Sauce Paradigm is an

aggression measure, then provoked participants should allocate a higher amount of the hot sauce than non-provoked participants in conditions without choice alternatives. However, they should not allocate more of the neutral or pleasant juice compared to non-provoked participants in conditions without choice alternatives. On the other hand, if the allocated amount of hot sauce is not just a mere measure of aggression but is also influenced by factors other than intentions to harm, a different pattern of results would emerge. Due to provoked participants' heightened responsiveness to the paradigm's salient cues (demand) in the direction of applying more of the provided option resulting, for example, from lowered self-control caused by rumination about an anger-inducing event (Denson, Pedersen, Friese, Hahm, & Roberts, 2011; Hirsh et al., 2011) provoked participants should allocate a larger amount of the aggressive option but also of the neutral and pleasant ones in comparison to non-provoked participants if they are not provided with choice options. With choice options, though, and thus with a lowered demand character of the paradigm, provoked participants should choose the aggressive option more often than non-provoked ones. However, since previous research with the Hot Sauce Paradigm might have overestimated the amount of aggression because the measured amount might not only reflect aggression but also responsiveness to demand, this difference in choices of provoked and unprovoked participants might not be significant.

3.1 Method

Participants and design

Participants were 152 students who participated for course credit or a candy bar. Nine participants expressed suspicion about the study's purpose and cover story or did not follow the instructions correctly (1 participant did not follow the instruction of not mixing the juices, 8 participants doubted that the confederate was an actual

participant). Their data was excluded from all further analyses leaving a total of 143 participants (48 male, unidentified 2). The design of the study was a 2 (provocation: no provocation vs. provocation) x 4 (choice options: single pleasant juice vs. single neutral juice vs. single spicy juice vs. all three juices to choose from) between-subjects design. As dependent variable I measured the juice choices as well as the allocated amount of the (chosen) juice.

Material and Procedure

Choice Options

I conducted a pretest with 19 students who participated voluntarily or for course credit to select the studies' stimuli. Instead of sauces, I decided to use juices as stimuli. With juice it was possible to provide options that would not be experienced as unpleasant even if larger amounts would have to be consumed. With sauces as stimuli this is questionable. If someone has to consume, for example, a cup of initially pleasant tasting chocolate sauce it might no longer be experienced as pleasant after consuming a certain amount of the sauce. I instructed participants to taste nine juice samples that included 3 potentially neutral, pleasant and spicy juices. To obtain spicy juices, I mixed regular juices with Tabasco sauce. After tasting a juice, participants rated the sample on three 9-point scales measuring how much they liked the juice's taste, how unpleasant it would be to consume a glass full of the juice, and how spicy the juice was. As pleasant juice stimulus I chose a multi-vitamin juice with the highest average rating for pleasant taste ($M = 8.11$, $SD = 1.41$), for the neutral juice stimulus I chose a fruit tea with a mid-scale rating indicating neither pleasant nor unpleasant taste ($M = 4.47$, $SD = 2.01$), and for the spicy juice stimulus I chose a juice that consisted of a mixture of orange juice (200 ml) and Tabasco sauce (2 teaspoons of

green Tabasco sauce) with a low rating for pleasant taste ($M = 1.79$, $SD = 0.85$) in combination with the highest rating for spicy taste ($M = 7.47$, $SD = 1.31$).¹

Affect measure

I measured participants' affect at three time points during the study. For the first time participants indicated their affective state after they completed a general taste preference questionnaire (baseline measure). For the second time I measured participants' momentary affect after they consumed a juice sample allegedly chosen from another participant that tasted either neutral (no provocation) or very sour (provocation). Finally, I measured participants' affective state after they had just prepared a juice sample for the target who ostensibly had selected the juice sample for them. To increase the feasibility of multiple measurement time points, I did not use all 60 items of the Positive and Negative Affect Schedule – Expanded Form (PANAS-X; Watson & Clark, 1994) previously used by McGregor et al. (1998). I chose the following six items to measure the affective state, with each representing the highest loading item for one dimension of the PANAS-X: *scared* (negative dimension: fear), *hostile* (negative dimension: hostility), *dissatisfied with the self* (negative dimension: guilt), *alone* (negative dimension: sadness), *concentrated* (positive dimension: attentiveness), *fearless*² (positive dimension: self-assurance), *happy* (positive dimension: joviality). To achieve a higher sensibility for changes across different measurement times I used visual analog scales with the poles *not at all* and *extremely* and instructed participants to place a cross on a 10 cm line indicating how they felt at the moment (Fahrenberg, 2006).

¹ The pleasant juice stimuli was “Milder Multivitamin” from the label “hohes C”, the neutral juice was the fruit tea of the store brand of the supermarket “Kaufland” (2 teabags in 1 liter of water), and the spicy juice was a mixture between the store brand orange juice of the supermarket “Kaufland” (200 ml) and 2 teaspoons of the green Tabasco sauce from the brand McIlhennys.

² Instead of using the highest loading item “bold” I used the item “fearless” with the second largest loading because the German translation of bold (German: *kühn*) is a rather uncommon word.

Procedure

The procedure of Study 1 was based on McGregor et al.'s (1998) Study 4 without the additional mortality salience manipulation. All other elements – expect that I measured the affective state three times – were modeled after McGregor et al.'s (1998) study. Participants that arrived at the lab were told that they would participate in a study that analyzes the effect of personality on taste preferences. After they signed a consent form, participants filled out a questionnaire about their general taste preference to enhance the credibility of the study's cover story. They also filled out the affect measures for the first time. Then the experimenter returned and participants received a covered cup with a juice sample ostensibly selected by another participant in an adjacent room out of a variety of juices with flavors ranging from neutral to very sour taste. All participants received a purple drink that either consisted of colored water (no provocation) or grape juice with vinegar (provocation). After consuming the entire juice sample, participants were instructed to rate the taste of the juice on a 9-point scale with the endpoints *neutral* (1) and *very sour* (9) as well as their liking of the juice on a 9-point scale with the endpoints *no liking at all* (1) and *extreme liking* (9). Then participants filled out the affect measure for the second time.

After a few minutes the experimenter returned with the experimental stimuli assembled on a tray. In the single juice conditions without choice options, the material contained one juice in a large cup (200ml; either the pleasant, neutral or spicy juice depending on condition) as well as a small cup (20ml) containing a taste sample of the same juice. In the choice option conditions, participants received three cups (200ml each) containing the pleasant, the neutral and the spicy juice as well as three small cups (20ml each) that contained taste samples of all three juices. On the tray the experimenter in all conditions also carried a white paper cup, aluminum foil to close

the cups and a glass of water. In addition, the experimenter gave participants a bogus taste preference questionnaire indicating the target's extreme distaste for spicy food. In the choice option conditions, participants then were instructed to try all three juices, select one of them, and administer an amount of their choice of this juice in the cup that the other participant from whom they just received their juice sample ostensibly would get. In the single juice conditions, participants were instructed to try the juice on the tray and administer an amount of their choice of this juice. In all conditions participants were informed that the target would have to consume the entire juice sample. After allocating the juice sample to the target, participants filled out a final questionnaire including the affect measure for the third time as well as control questions used by McGregor et al. (1998) asking participants to rate the extent to which they used the taste preference inventor as basis for their allocation on a 21-point scale with the endpoints *not at all* (1) and *completely* (21), the usefulness of the taste preference questionnaire on a 21-point scale with the endpoints *not at all useful* (1) and *extremely useful* (21), and how much they thought the other alleged participant would like the allocated juice sample on a 9-point scale with the endpoints *no liking at all* (1) and *extreme liking* (9). In addition, I asked participants to rate their own liking of the allocated juice sample on a 9-point scale with endpoints *no liking at all* (1) and *extreme liking* (9) and the likeability of the other alleged participant with a 5-point scale from *very likable* (1) to *very unlikable* (5) as well as, in the conditions with choice options, an open-ended question asking participants to indicate the strategy they used to choose the juice and, in all conditions, an open-ended question about the possible purpose of the study. After answering these control questions, participants filled out a German version of the 12-item short version of the Buss-Perry Aggression Questionnaire (Bryant & Smith, 2001) and answered demographic questions (age, sex, profession, major if they indicated student as profession, and first

language). Finally, the experimenter returned and thoroughly debriefed the participants. I especially took care to ensure that none of the participants left with any kind of negative feeling due to the participation in a study that included drinking a sample and possibly allegedly allocating an unpleasant tasting juice to someone else. It was made clear to participants that nobody had to drink the juice sample they prepared. They were also assured that their allocation did not reflect any negative aspects of their personality but that they responded to situational variables of the study designed to measure certain constructs. No participant indicated any form of distress or voiced negative feelings. In contrast, many participants mentioned that they liked participating in the study especially since it was not just a paper-pencil or computer-administered study.

3.2 Results

Juice choices

I assessed the juice choices in the two experimental groups with choice options as one dependent variable. With choice options, only three of the 18 provoked participants chose the neutral or hot juice (1 neutral, 2 spicy juice choices) and none of the 20 non-provoked ones did, $\chi^2(2) = 3.62, p = .16$. All three participants who did not choose the pleasant juice indicated in the open-ended question asking them about their choice strategy that they specifically chose the spicy or neutral juice because of its particularly bad or at least not pleasant taste, respectively. For the pleasant juice choices participants most commonly listed as strategy that they chose the juice that tasted the best for themselves and thus most likely would be liked by the target as well.

Juice allocation

Comparable to the dependent variable in the classical Hot Sauce Paradigm, I collected the weight in gram of the juice allocated to the target person as dependent variable. On average male participants allocated more juice, $M = 77.27$, $SD = 43.79$, than female participants, $M = 66.58$, $SD = 38.50$. However, this difference was not significant, $t(139) = 1.49$, $p > .10$. Thus, in all further analyses data for both genders were collapsed and analyzed together.

First, I compared the two experimental groups that could only administer different amounts of the spicy juice and that were either provoked or not to test if I replicated previous findings with the Hot Sauce Paradigm. Considering only these two groups, the present study resembles the classical Hot Sauce Paradigm studies using juice instead of sauce. A t -test indicated an effect of provocation on the allocated amount of the spicy juice, $t(34) = 1.78$, $p < .05$, one-tailed, replicating previous findings. A one-tailed test was used since the direction of the effect was expected based on these previous findings. Provoked participants allocated more spicy juice, $M = 46.00$, $SD = 42.17$, than non-provoked ones, $M = 26.41$, $SD = 17.32$.

A 2 (provocation vs. no provocation) x 4 (single alternative pleasant juice vs. single alternative neutral juice vs. single alternative spicy juice vs. all three juices to choose from) ANOVA on the allocated amount of juice revealed a significant main effect of provocation, $F(1, 135) = 7.50$, $p < .01$, with provoked participants allocating more juice than non-provoked ones, as well as a main effect for the available type of juice, $F(3, 135) = 20.53$, $p < .01$. Post-hoc Scheffe's Tests indicated that participants allocated significantly less of the single alternative spicy juice compared to any other single alternative juice or chosen juice in the condition with choice options (mean difference in allocated amount between spicy and pleasant juice: $M_{diff} = 58.31$, $p <$

.05, between spicy and neutral juice: $M_{\text{diff}} = 28.88$, $p < .05$, and between spicy and the chosen juice with choice options: $M_{\text{diff}} = 48.98$, $p < .05$). They also allocated significantly less of the single alternative neutral juice than of the single alternative pleasant one, $M_{\text{diff}} = 29.25$, $p < .05$.

Individual analyses indicated that provoked participants not only administered more of the spicy juice, $F(1, 142) = 2.94$, $p < .05$, one-tailed (direction of the effect was expected based on previous research), but they also administered more of the pleasant juice, $F(1, 142) = 7.16$, $p < .01$. The descriptive pattern for the neutral juice was in line with predictions but the difference did not reach conventional significance levels, $F(1, 142) = .47$, $p > .10$. A combined t -test of the pleasant and neutral juice without choice options indicated a significant difference in the allocated juice, $t(67) = 2.32$, $p < .05$. When participants were provided with choice options, there was no significant difference in the chosen amount between provoked and non-provoked participants, $F(1, 142) = .12$, $p > .10$ (see Table 1).

Table 1

Means and standard deviations for allocated amount of juice (in grams) for provoked and non-provoked participants with or without choice options

Choice Options	Provocation	Allocated juice		
		<i>N</i>	<i>M</i>	<i>SD</i>
Only spicy juice	No provocation	17	26.41	17.32
	Provocation	19	46.00	42.17
Only neutral juice	No provocation	18	61.78	38.06
	Provocation	17	69.71	28.07
Only pleasant juice	No provocation	17	79.18	33.22
	Provocation	17	110.60	29.38
All three juices to choose from	No provocation	20	83.90	38.89
	Provocation	18	87.78	37.44

Affect measures.

I measured participants' affective state at 3 time points: as a baseline measure (T1), right after the provocation (T2) and after the participants allocated the juice sample to the target (T3). With measures from these 3 time points and the additional between-subject factors (provocation vs. no provocation and single alternative pleasant juice vs. single alternative neutral juice vs. single alternative spicy juice vs. all three juices to choose from) as well as the interaction between these two factors I conducted repeated-measures ANOVAs for each of the seven affect scales.³ The analyses resulted in a marginally significant Time x Provocation interaction for the participants' hostility level, $F(2, 278) = 2.66, p = .07$ (see Table 2). All other interaction effects between the provocation and the choice option factor with measurement time were non-significant. To explore the interaction between measurement time and provocation on hostility further I conducted separate ANOVAs for each measurement point. Detailed analyses revealed that there was no significant difference in the hostility level between provoked and non-provoked participants (all $p > .10$) except at T1 with a marginally significant higher hostility level in the no provocation groups, $F(1, 140) = 3.12, p = .08$. Simple analyses including only provoked participants revealed a significant effect of the measurement time, $F(2, 136) = 5.13, p < .01$. Separate paired t -test resulted in a significant increase in hostility right after the provocation T1 and T2, $t(69) = -2.95, p < .01$, and a significant decrease after the participants allocated the juice to the target between T2 and T3, $t(69) = 1.99, p = .05$. The difference between T1 and T3 was not significant, $t(69) = -$

³Before conducting the analyses variables were inspected for outliers. Outliers were defined as cases with standardized z -values larger than 3.29 or smaller than -3.29 that appeared disconnected from the other cases (Tabachnick & Fidell, 2007). In total, on all affective state measures across all three time points 16 outliers were identified. As recommended by Tabachnick and Fidell (2007), variable values in these cases were replaced by a value that was half a scale point higher than the rating in the next extreme case.

1.43, $p > .10$. There also was a marginally significant change in hostility in a separate analyses including only non-provoked participants, $F(1, 142) = 2.50$, $p > .09$. Separate paired t -test resulted in no significant difference in hostility between T1 and T2, $t(71) = -0.53$, $p > .01$, but a significant decrease after the participants allocated the juice to the target between T2 and T3, $t(69) = 2.11$, $p < .05$. Also, there was a marginally significant increase in hostility between T1 and T3, $t(71) = 1.74$, $p = .09$.

Trait Aggression

In a correlation analysis with the self-reported overall trait aggression as well as the physical aggression subscale and allocated juice in grams I tested the relation between juice allocation and trait aggression.⁴ Ignoring the allocated type of juice, there was neither a correlation between the allocated amount of juice and overall trait aggression, $r = -.01$, $p > .10$, $n = 138$ due to missing value, nor physical aggression, $r = .08$, $p > .10$, $n = 142$ due to one missing value. In a separate analysis, in which I only considered the allocated amount of the spicy juice in the no choice option conditions, I also found no correlation between overall trait aggression, $r = -.12$, $p > .10$, $n = 35$ due to one missing value, or physical aggression, $r = .07$, $p > .10$, $n = 36$, and allocated amount of spicy juice.

⁴Before conducting the analysis the variables were inspected for outliers. Again, outliers were defined as cases with standardized z -values larger than 3.29 or smaller than -3.29 that appeared disconnected from the other cases (Tabachnick & Fidell, 2007). Two outliers were identified and replaced by a value that was half a scale point higher than the rating in the next extreme case.

Table 2

Means and standard deviations for participants' hostility level at measurement times T1, T2, and T3

		<i>N</i> ⁵	<i>M</i>	<i>SD</i>
No provocation	T1	72	1.12	1.13
	T2	72	1.17	1.23
	T3	72	0.96	0.97
Provocation	T1	70	.82	.83
	T2	70	1.21	1.31
	T3	71	.95	1.00

3.3 Discussion

Using a modified Hot Sauce Paradigm in the first study, I replicated the results of previous research in that provoked participants administered significantly more spicy juice when they were provided with a single spicy juice only and no other choice options. However, I also demonstrated a main effect of provocation on the allocated amount of juice. Provoked participants administered significantly more juice than non-provoked ones indicating that provoking participants not only affected the potential aggressive behavior but also other variables related to the allocated amount of juice. In other words, the fact that provocation also increased the allocated amount of non-aversive juices suggests that increases in the allocated amounts of juice might be driven by factors other than intentions to harm the other person. A possible explanation might be that the effect is caused by a lowered self-control in the provocation conditions due to rumination about the anger inducing provocation (Denson et al., 2011) and, with that, the reactivity to demand in the direction of applying more juice inherent in the paradigm might be increased (Hirsh et al., 2011).

⁵ At T1 and T2, 1 provoked participant per measurement time failed to provide an answer resulting in 1 missing value per question,

This possible explanation was further analyzed with Study 2 and 3. No matter what process causes the effect, the results of Study 1 question the validity of the allocated hot sauce as an aggression measure.

Study 1 also highlighted the importance of providing choice options to avoid underestimating the amount of non-aggressive behavior shown by participants. With choice options provoked participants rarely chose a juice other than the pleasant one. A reason for the low frequency of aggressive choices might be that the applied provocation was not very intense. To provoke participants I replicated the method previously used by McGregor et al. (1998) and only made them drink a sour tasting juice sample but they were never personally offended or in any other form provoked on an interpersonal level.

Unlike McGregor et al. (1998) who measured participants' affective state only once I explored the development of affective states over the course of the Hot Sauce Paradigm in the first study. The results indicated changes in participants' hostility level with an increase right after the provocation at T2 and a decrease in the hostility level at T3 for provoked participants. The increase in hostility might be the result of an effective provocation. The decrease in hostility after participants interacted with the target that provoked them beforehand by allocating a juice to that person might be the result of some form of goal fulfillment. This might be the fulfillment of a revenge goal, a goal to restore justice or just a goal to achieve an equilibrium by making the target drink something too. Goal fulfillment then might cause a lowered accessibility of aggressive and hostile thoughts (Denzler, Förster, & Liberman, 2009) which might also explain the decrease in hostility between T2 and T3 for non-provoked participants. Yet, there was no significant difference in hostility of provoked and non-provoked participants right after the provocation at T2 but a marginally significant

difference at T1 with participants in the no provocation conditions indicating slightly more hostility at the baseline measure as well as a marginally significant increase in hostility between T1 and T3 in the no-provocation condition. I do not have an explanation for this difference; it might just be a random effect. It is necessary to further analyze changes in regard to the affective state during the Hot Sauce Paradigm. The results of this study should be taken into account cautiously considering the very low level of hostility throughout the entire study with means around 1 on 10-point scales (see Table 2) and, thus, possible bottom effects. In addition, some participants uttered their confusion about having to answer the same questions more than once, which might make it especially hard to detect small changes in affective states. Hence, it might be worth it to investigate changes in affective states during the paradigm with a different, more intense form of provocation with which larger changes can be expected.

In regard to trait aggression, I was unable to replicate the correlation between general as well as physical aggression and the amount of allocated hot juice. However, I did not control for the moderator of relational versus experiential thinking mode identified by Liebermann et al. (1999). The evidence for a significant correlation between the allocated sauce amount and physical aggression provided by these authors was limited to the rational thinking mode characterized by more deliberate thinking. In an experiential thinking mode characterized by a quick, intuitive driven thinking style there was no correlation. Furthermore, since I used a short version of the Buss Perry Aggression questionnaire (Bryant & Smith, 2001) the trait aggression measure might be less sensitive than the regular length scale. To analyze the relation with trait aggression further in the following studies I used the

regular 27-item German version of the Aggression Questionnaire instead (Herzberg, 2003).

4 Study 2

In Study 1 I found that provoked participants applied more of a given juice than non-provoked ones. With choice options this main effect was no longer significant. With the instruction used in the Hot Sauce Paradigm in which participants are asked “to put as much or as little” (Lieberman et al., 1996, p. 339) sauce in the cup for the alleged other participant demand is created by specifically focusing on the amount of allocated juice. This might guide participants’ behavior toward someone especially when this person has just provoked them in an uncommon situation. It also cues using the allocated amount of juice as a way to handle the situation, which might not be a strategy that participants would come up with spontaneously. In addition, the instruction contains permissive cues for applying large amounts of juice to the target (Ritter & Eslea, 2005). Provoked participants might be especially likely to follow leading instruction due to a lowered self-control caused by rumination about an anger-inducing event (Denson et al., 2011).

I formulated a less demand-prone version of the instructions in addition to the one used in Study 1 and contrasted both in a setting that replicated the single alternative spicy juice condition of Study 1. I hypothesized to find a difference between the allocated juice amount of non-provoked and provoked participants with the regular instruction as used in Study 1 with provoked participants applying more juice than unprovoked ones. No such difference was expected between the allocated juice amount of non-provoked and provoked participants with the less demand-prone instruction. In addition, I measured participants’ state self-control level right after the

provocation to test whether a difference in self-control could explain the predicted difference in the regular instruction conditions.

4.1 Method

Participants and design

Participants were 90 students who participated for course credit or a candy bar. 14 participants expressed suspicion about the study's purpose and cover story or did not follow the instructions correctly (2 participants refused to drink the unpleasant juice, 3 participants mixed the juice with the provided water, 3 participants doubted that the confederate was an actual participant, 6 participants conjectured that the study's true purpose might in a broader sense be revenge behavior). Their data was excluded from all further analyses leaving a total of 76 participants (16 male, 8 unidentified). The study followed a 2 (provocation: no provocation vs. provocation) x 2 (instruction: regular instruction vs. less demand-prone instruction) between-subjects design. As a dependent variable I measured the allocated amount of spicy juice.

Material and procedure

The procedure of Study 2 was identical to the single alternative spicy juice conditions in Study 1 with the following exceptions. Instructions in the less demand-prone instruction condition were a reformulated version of the instructions used in Study 1. More precisely, I reformulated the instruction given to participants before they prepare the juice sample for the other alleged participant by removing phrases specifically addressing the amount of juice and referring to it more generally as juice sample and thus taking out cues that potentially guide participants in regard to how they should behave in an unfamiliar situation (Ritter & Eslea, 2005; full instructions as used in Study 2 see Appendix 1). Also, I did not measure the affective state in this study. Instead, to measure the impact of the provocation manipulation on participants'

self-control, participants filled out the 10-item State Self-Control Capacity Scale (Twenge, Muraven, & Tice, 2004) right after they received the juice sample from the other alleged participant (Cronbach's $\alpha = .80$).

After participants allocated the amount of juice and filled out an open-ended question about the studies purpose⁶ they filled out the German translation of the Buss-Perry Aggression Questionnaire (Herzberg, 2003) and answered demographic questions (age, sex, profession, major if they indicated student as profession, and first language). Then the experimenter returned and participants were thoroughly debriefed as described in Study 1.

4.2 Results

Juice allocation

To test the relation between provocation and juice allocation with the regular and the less demand-prone instructions I performed a *t*-test and an ANOVA with the allocated juice as dependent variable.⁷

⁶To get further insights in the validity of the paradigm, I attempted to measure participants' harmful intentions, as one key concept for the definition of aggression (Tedeschi & Felson, 1995; Ritter & Eslea, 2005) using a questionnaire that participants filled out before answering questions referring to trait aggression. Based on Tedeschi and Felson's (1995) definition of intentional actions 4 items were formulated to measure the harmful intention behind the amount allocation: "What do you think, did the other participant like the drink that you have allocated?" 5-point scales with poles *no liking at all* (1) and *extreme liking* (5); "How unpleasant do you think it was for the other participant to drink the entire sample of the drink that you allocated?" 5-point scales with poles *very unpleasant* (1) and *very pleasant* (5); "How much would you like it if the other participant would not at all like the drink?" 5-point scales with poles *no liking at all* (1) and *extreme liking* (5); "How much would you like it of the other participant really like the drink?" 5-point scales with poles *no liking at all* (1) and *extreme liking* (5). Items 1, 2 and 4 were reversed scored. Thus, higher values reflected more harmful intentions. However, the inter-correlation and the Cronbach's α was around zero ($\alpha = -.01$). Thus, I refrained from further discussing the results related to the intention measure. Participants might have had problems answering questions for which awareness about the allocated amount was necessary.

⁷Before conducting the analysis the variables were inspected for outliers. Again, outliers were defined as cases with standardized *z*-values larger than 3.29 or smaller than -3.29 that appeared disconnected from the other cases (Tabachnick & Fidell, 2007). Three outliers were identified for the allocated amount of juice. Since, these extreme cases might be due to a misunderstanding of the instruction (e.g., the participants allocated the entire amount of the available juice possibly because they did not understand that they could chose the amount) these three cases were excluded from the analyses paralleling the strategy to handle outlier in the allocated juice used by Lieberman et al. (1999) thus leaving $n = 73$ participants (14 males, 8 unidentified).

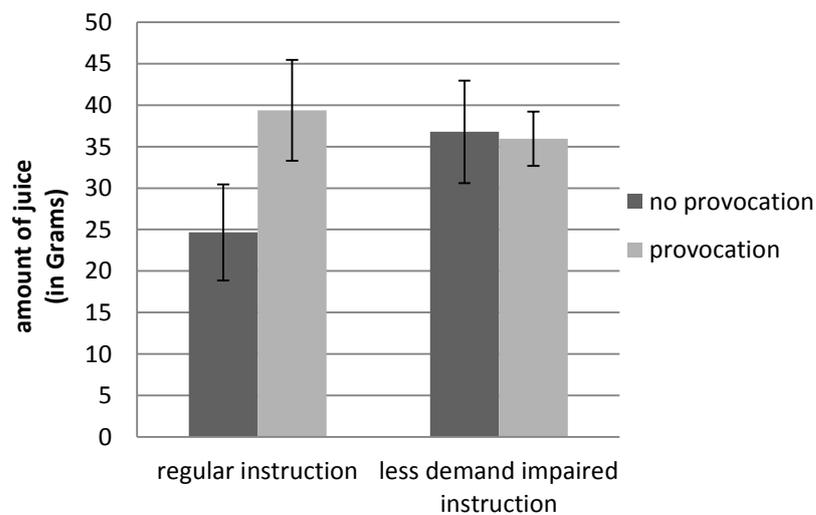
On average male participants allocated more juice, $M = 41.71$, $SD = 19.03$, than female participants, $M = 33.59$, $SD = 25.39$. However, this difference was not significant, $t(63) = 1.11$, $p > .10$. Thus, in all further analyses data for both genders were collapsed and analyzed together.

To check if I was able to replicate the findings of Study 1, I compared the juice allocation of provoked and non-provoked participants for the condition in which data was collected with the same instruction as in Study 1. A t -test revealed an effect of provocation on the allocated amount of the spicy juice, $t(34) = 1.76$, $p < .05$, one-tailed (since the direction of the effect was specified in the hypotheses), replicating the findings of Study 1 and previously found with the original Hot Sauce Paradigm. Provoked participants allocated more juice, $M = 39.38$, $SD = 27.87$, than non-provoked ones, $M = 24.07$, $SD = 22.41$.

In the full data set provoked participants also allocated more juice ($M = 73.79$, $SD = 22.30$) than non-provoked ones ($M = 31.18$, $SD = 25.48$). However, a 2 (provocation vs. no provocation) \times 2 (regular instruction vs. less demand impaired instruction) ANOVA on the allocated amount of juice revealed that this difference did not reach conventional levels of significance, $F(1, 69) = 1.67$, $p = .20$. The same was the case for the interaction between provocation and type of instruction, $F(1, 69) = 2.08$, $p = .15$. Since I formulated specific hypotheses about the differences between the experimental groups, I nevertheless ran simple main effect analyses. The analyses revealed that provoked participants allocated significantly more juice than non-provoked participants when the data was collected with the regular instruction, $F(1, 69) = 3.64$, $p < .05$, one-tailed (since the direction of the effect was specified in the hypotheses). However, there was no significant difference in the juice allocation when the data was collected with the less demand impaired instruction, $F(1, 69) < 1$.

Simple main effect analyses also revealed that without provocation participants allocated marginally more juice with the less demand impaired instruction than with the regular one $F(1, 69) = 2.41, p = .07$. There was no such significant difference for provoked participants, $F(1, 69) < 1$ (see Figure 1).

Figure 1



Average allocated amount of juice for provoked and non-provoked participants when data was collected with the regular versus the less demand impaired instruction, *SE* as error bars

Self-control

To examine whether the interaction between the type of instruction and provocation on the allocated juice amount was due to a difference in self-control between provoked and non-provoked participants and, with that, potentially a higher responsiveness to the demand of the depleted participants I first conducted a *t*-test. With this test I examined whether I replicated previous findings of a decreased self-control after experiencing a provoking event (Denson et al., 2011). However, my analyses revealed no difference in self-control between provoked, $M = 5.44$, $SD = 0.74$, and non-provoked participants, $M = 5.41$, $SD = 1.11$, $t(72) < 1$. Thus, I refrained from performing further analyses with self-control in this study.

Trait aggression

To test the relation between juice allocation and trait aggression I performed a correlation analysis with the self-reported overall trait aggression as well as the physical aggression subscale and allocated juice in gram.⁸ There was no correlation between the allocated amount of juice and overall trait aggression, $r = -.01$, $p > .10$, or physical aggression, $r = -.14$, $p > .10$.

4.3 Discussion

I replicated the allocation results gained with the Hot Sauce Paradigm in previous studies and in Study 1 with provoked participants allocating more spicy juice than non-provoked ones. However, this effect disappeared in the condition in which data was collected with less guiding instructions. Here, I found no difference between provoked and non-provoked participants in regard to their juice allocation. As a

⁸Before conducting the analysis the total aggression and physical aggression score were inspected for outliers. Again, outliers were defined as cases with standardized *z*-values larger than 3.29 or smaller than -3.29 that appeared disconnected from the other cases (Tabachnick & Fidell, 2007). Two outliers were identified and replaced by a value that was half a scale point higher than the rating in the next extreme case.

potential reason for the difference in the regular instruction condition I proposed a difference in self-control between provoked and non-provoked participants. Due to this difference, I assumed that provoked participants would follow guiding instructions more than non-provoked ones. However, I did not find a difference in self-control between provoked and non-provoked participants and thus was unable to replicate Denson et al.'s (2011) results making the proposed self-control explanation for the observed difference in the allocation less likely. This failure to replicate previous research might be due to a different, less intense provocation method since Denson and colleagues (2011) used a more severe interpersonal provocation in which participants in the provocation condition received insulting feedback allegedly given by another participant. Also possible is that the self-control questionnaire (Twenge et al., 2004) was not sensitive enough to capture small changes in self-control. However, Denson et al. (2011) documented changes in self-control using this scale. Even though the data of Study 2 questioned the self-control explanation, I attempted to replicate the effect of provocation on self-control for a second time with Study 3 combining it with an experimental manipulation of the easiness of the juice application.

For trait aggression, I again was not able to replicate the correlation between physical aggression and the amount of allocated hot juice using the regular 27-item German version of the Aggression Questionnaire (Herzberg, 2003) instead of using the 12-item short version (Bryant & Smith, 2001) used in Study 1. Thus, the lack of correlation between trait aggression and the juice amount in Study 1 was most likely not due to a potentially less accurate trait measure.

5 Study 3

With the third study I again attempted to replicate the effect of provocation on self-control. Instead of using more or less demand-prone instructions, I orthogonally crossed the provocation manipulation used in Study 1 and 2 with a manipulation making the application of the juice more or less difficult and thus making more or less self-control necessary to follow the demand in the direction of applying more juice. Otherwise the procedure was similar to the regular instruction condition of Study 2.

I hypothesized that without provocation the difficulty in the application of the juice would lead to a smaller difference between the easy and hard application groups than with provocation. I assumed that without provocation participants would have the necessary self-control that would enable them to apply just as much juice as they intended independent of the easiness of the application. In the provocation condition this should not be given to the same extent with provoked participants applying more or less juice depending on the easiness of the application due to a difference in self-control.

5.1 Method

Participants and design

Participants were 88 students who participated for course credit or a candy bar. 11 participants expressed suspicion about the study's purpose or the cover story or did not follow the instructions (1 participant was familiar with the Hot Sauce Paradigm, 1 participant doubted that the confederate was an actual participant, 2 participants refused to drink the juice sample, 7 participants conjectured that the study's true purpose might in a broader sense be revenge or anti-social behavior). Their data was excluded from all further analyses leaving a total of 77 participants (33 male, 6 unidentified). The study followed a 2 (provocation: no provocation vs.

provocation) x 2 (application: easy vs. hard) between-subjects design. As a dependent variable the allocated amount of spicy juice was measured.

Material and procedure

With the following exceptions the procedure of the Study 3 was identical to the regular instruction condition in Study 2.⁹ In addition to the provocation manipulation I orthogonally manipulated the easiness of applying a large amount of juice in the cup for the other participant. When the experimenter returned after participants consumed the sour (provocation) or neutral juice sample (no provocation) he/she carried a small bottle (200 ml) filled with the spicy juice instead of a large cup filled with this juice on a tray. The bottle was specially prepared so that the bottle's cap either had a very small or wide opening. A pretest with $n = 20$ students revealed that participants experienced the juice application significantly less easy, $t(18) = 3.66$, $p < .01$, with the small opening in comparison to the large one. Thus, the application of a large amount of juice was either hard or easy depending on which condition participants were assigned to. In addition, the bottle was covered with tinfoil, so that the experimenter could not see how much juice participants allocated to the other alleged participant. Though I replicated the provocation method used by McGregor et al. (1998) I also asked participants to fill out 2 questions in this study with which I intended to check whether or not the chosen provocation method was

⁹ Like in Study 2 I attempted to measure participants' harmful intentions with the same 4 items that participants filled out before answering questions referring to trait aggression. Analyzed separately for easy and hard to allocate juice the internal consistency of the items reached values that might justify computing an average intention measure over the 4 items, $\alpha_{\text{easy allocation}} = .56$, $\alpha_{\text{hard allocation}} = .65$. However, for the easy allocation group the value was still rather low and might, again like in Study 2, reflect potential problems of the participants in regard to answering questions for which an awareness of the allocated amount was necessary. There was a significant correlation between harmful intentions and the allocated amount in the hard to allocate group, $r = .42$, $n = 39$, $p < .01$. For the easy allocation group this correlation was not significant, $r = .13$, $n = 37$, $p > .10$. Since the main focus of the application easiness manipulation was the manipulation of the easiness of applying a specific amount of juice and because of the rather low internal consistency as well as the inconsistent results over the two experimental groups, the results should be regarded carefully and an interpretation of these results might be problematic. Thus, I refrained from further discussing them in the main section.

successful and how intensely it was experienced by participants. Before filling out the trait aggression questions I asked them to rate the likability of the other alleged participant on a 5-point scale with the endpoints *very likable* (1) and *very unlikable* (5) as well as how much they felt provoked by the other participants behavior on a 5-point scale with the endpoints *not provoked at all* (1) and *extremely provoked* (5). Then the experimenter returned and participants were thoroughly debriefed as described in Study 1.

5.2 Results

Manipulation check

In the provocation conditions, participants felt more provoked than in the no provocation conditions, $t(72)^{10} = -2.35, p < .05$, Cohen's $d = .57$, and took less liking in the confederate without, however, reaching conventional levels of significance, $t(72)^{11} = -1.51, p = .14$, Cohen's $d = .37$. In sum this suggesting an effective provocation manipulation with small to medium effects.

Juice allocation

To test the relation between provocation and juice allocation with easy or hard to allocate juice I performed an ANOVA with the allocated juice as dependent variable.¹² There was no significant difference in the juice allocation of male, $M = 25.59, SD = 17.69$, and female participants, $M = 25.03, SD = 17.00, t(69) < 1$. Thus, in all further analyses data for both genders were collapsed and analyzed together.

A 2 (provocation vs. no provocation) x 2 (easy application vs. hard application) ANOVA on the allocated amount of juice revealed that participants

¹⁰ 4 participants failed to provide an answer to this question resulting in 3 missing values, $n = 74$

¹¹ 4 participants failed to provide an answer to this question resulting in 3 missing values, $n = 74$

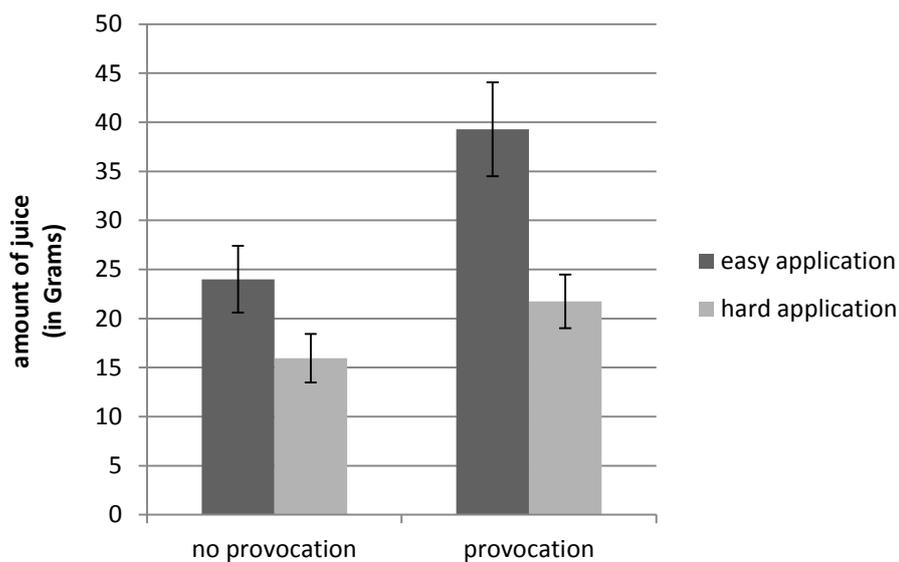
¹² Before conducting the analysis the variable allocated amount of juice was inspected for outliers. Again, outliers were defined as cases with standardized z -values larger than 3.29 or smaller than -3.29 that appeared disconnected from the other cases (Tabachnick & Fidell, 2007). One outliers was identified and excluded from the analyses like in Lieberman et al. (1999) thus leaving $n = 76$ participants.

allocated significantly more juice when the allocation was easy in comparison to when it was hard, $F(1, 72) = 14.09, p < .01$. Provoked participants also allocated more juice than non-provoked ones, $F(1, 72) = 9.54, p < .01$. However, the predicted interaction between allocation easiness and provocation was not significant, $F(1, 72) = 1.94, p = .17$. The descriptive pattern indicated that the interaction was not in the predicted direction. For both easy and hard allocation conditions provoked participants allocated more juice than non-provoked ones (see Figure 2).

Self-control

To examine whether I replicated previous findings of a decreased self-control after provocation (Denson et al., 2011), I conducted a t -test on the self-control of provoked and non-provoked participants. However, my analyses again revealed no difference in self-control between provoked, $M = 5.33, SD = 0.83$, and non-provoked participants, $M = 5.34, SD = 1.19, t(72) < 1$.

Figure 2



Average allocated amount of juice for provoked and non-provoked participants for easy to allocate versus hard to allocate juice; SE as error bars

5.3 Discussion

With the third study I again attempted to test the assumption of higher responsiveness to the paradigm's demand due to lowered self-control after provocation as potential explanation for the main effect on the amount of allocated juice found in Study 1. I manipulated the easiness to apply more juice and thus to follow the paradigm's demand. However, I did not find a significant interaction between allocation easiness and provocation. I also was unable to replicate a difference in self-control between provoked and non-provoked participants making the self-control explanation rather unlikely for the juice type independent main effect of provocation on the applied amount in Study 1. A possible alternative might be that provoked participants ruminate about the provoking event and thus are less focused on the study. This could heighten the likelihood for them to follow a demand and thus apply more juice than non-provoked participants but without being intense enough to actually lower self-control. In comparison to the provocation method chosen by Denson et al. (2011) the applied method did not include any form of personal insult and thus might be a less intense one. In Study 1, wherein I used the same provocation method as in 2 and 3, the amount of aggressive juice choices in the choice option condition was low possibly due to a rather low intensity of provocation. In addition, the effect sizes on the manipulation check questions used in this study revealed small to medium effects of the provocation manipulation on experienced provocation and subjective likability of the alleged other participant. Thus, in the fourth study I used a more intense form of provocation and measured state self-control after the provocation.

6 Study 4

The major focus of the fourth study was to further analyze the choice-based modified version of the Hot Sauce Paradigm. My goal was to examine whether this modified paradigm might be useful for the application in aggression research. In Study 1 I found a relatively low occurrence of behavioral aggression. In conditions in which participants were provided with choice options only 3 of 18 provoked participants choose a non-pleasant (neutral or spicy) juice for the target. The possibility of analyzing aggressive behavior and the conditions under which it occurs with a specific paradigm are severely limited if the behavior the research focus lies upon is only shown by such a small number of participants. Due to the low number of respective choices, I was unable to, for example, analyze the relation between juice choices and trait aggression or intentions in Study 1. The low choice rates of the aggressive option might be due to the use of a rather low intensity provocation method in this study. The small to medium effect sizes found on the manipulation check questions in Study 3, in which I applied the same provocation method, makes this explanation likely. Thus, with the fourth study I intensified the provocation in order to analyze whether this would lead to more neutral or aggressive choices. With a higher amount of spicy or neutral juice choices I would be able to analyze the validity of the choice-based paradigm further. Especially, it would enable me to analyze the relation between the juice choices and central, aggression defining constructs like harmful intentions.

I switched to an interpersonal provocation to intensify the provocation. Interpersonal provocations are known as one of the most intense and powerful methods for eliciting aggression (Bettencourt & Miller, 1996; Carlson & Miller, 1998) and are regarded as a key cause of human aggression (Anderson & Bushman,

2002). Specifically, I combined paradigms previously used by Holmes (1972) and Atkinson and Polivy (1976). More precisely, in the provocation condition participants either waited for 15 minutes for another alleged participant that behaved rude toward them when he/she arrived at the laboratory. In the no provocation condition they only waited for a short moment and the other alleged participant showed no rude behavior. Since waiting alone does not necessarily lead to anger and aggression against the person that is late I combined rude behavior with waiting time (Atkinson & Polivy, 1976).¹³

In some interpersonal provocation paradigms used in previous research participants see the target that provokes them. Participants, for example, meet a confederate in a waiting room who is either late (provocation) or on time (no provocation) for an alleged mutual study (Holmes, 1972), are verbally attacked by (Atkinson & Polivy, 1976) or receive negative feedback from the experimenter (Caprara, Renzi, Alcini, D'Imperio, & Travaglia, 1983) or receive a negative personality rating by a confederate that they see before the actual experiment starts (Baron, 1979). Opposed to this, one shortcoming of laboratory aggression research is a large unvaried distance between the aggressor and the victim. In many studies participants are informed that the victim supposedly is in an adjacent room and thus the victim is at no point visible for participants. In the studies using and interpersonal provocation, in which participants see the target, the victim's visibility is not varied between participants and thus potential effects of victim visibility on aggressive behavior are unclear. To analyze how victim visibility affects participants' behavior, I orthogonally manipulated this variable with the provocation factor. I expect a lowered amount of aggression with visible in comparison to invisible victims potentially due

¹³ In an attempt to find a suitable provocation method, one study was started in which I tried to provoke participants only with a rude comment of the alleged other participant. However, I stopped the data collection after $n = 20$ participants since aggressive choices did not occur.

to emotional attachment to the victim based on Milgram's (1974) obedience research and replications of this research with the Teacher-Learner Paradigm (Ahmed, 1979; Page & Moss, 1976) as well as field studies with either visible or non-visible confederates serving as target victims after they showed rude behavior in traffic (Turner, Layton, & Simons, 1975). All these experiments documented lowered amounts of obedience and aggressive behavior for visible in comparison to invisible victims. The victim's human nature might be more salient with an image of the victim in mind and thus that might lead to less anonymous conditions and, consequently, to less aggressive behavior (Bandura, 1975; Diener, Fraser, Beaman, & Kelem, 1976).

In sum, I expected more aggressive choices in the provocation than in the no provocation condition and that these choices would be mediated by participants' harmful intentions. In addition, I expected fewer aggressive choices with visible in comparison to non-visible victims. Since all participants were provided with choice options, the demand of the modified paradigm is lower in comparison to the classical one. Thus, I did not expect to find an effect of provocation on the allocated amount of juice.

6.1 Method

Participants and design

Participants were 92 students who participated for course credit or a small payment (4 EUR, approximately 5 USD). 21 participants expressed suspicion about the study's purpose, the cover story or did not follow instructions correctly (1 participant did not follow the instruction of not mixing the juices, 1 participant was familiar with the Hot Sauce Paradigm, 6 participants doubted that the confederate was an actual participant, 13 participants conjectured that the study's true purpose might in a broader sense be revenge behavior). Their data was excluded from all further

analyses leaving a total of 71 participants (19 male). The study followed a 2 (provocation: no provocation vs. provocation) x 2 (visibility: not visible vs. visible victim) between-subjects design. As dependent variables I assessed the juice choices as well as the allocated amount of the chosen juice.

Material and Procedure

Confederates

To avoid effects caused by the person giving the instructions or playing the confederate, 12 research assistants served as confederates and 11 as experimenters. They were all psychology majors and intensely trained in how to play the role of the confederate or experimenter.

Intention measure

Intentional actions are defined by Tedeschi and Felson (1994) as actions that are performed with the expectation that they will lead to a proximate outcome. This outcome is causally linked to a terminal goal and, due to this relation, the outcome is valued. According to this definition, a person has harmful intentions if he or she performs an action with the expectation to harm someone and if this harm doing is valued positively by that person. On the basis of this definition, I formulated three items to measure the intention behind juice choices. To measure the extent of expectations to harm, participants filled out the following item “What do you think, did the other participant for whom you selected the drink like the drink that you have chosen?” on a 5-point scale with the poles *no liking at all* (1) and *extreme liking* (5). To measure how much participants valued the harm caused by their juice choices, participants answered the items “How much would you like it if the other participant would not at all like the drink that you have chosen?” and “How much would you like

it if the other participant would really like the drink that you have chosen?" on the same 5-point scale as the first item. I computed an average score (Cronbach's $\alpha = .76$) with reverse coding of the first and third item as a measure of harmful intentions. Higher values thus represent more harmful intention.

Procedure

To participate in the study participants scheduled an appointment either at the university cafeteria, in class or via an online registration on the university's web page. Participants were made believe that it was necessary to schedule an appointment because the study allegedly involved two people participating at the same time.

When participants arrived at the laboratory they were asked to hand their bags and cell phones to the experimenter to ensure that they would not distract themselves in case they had to wait. The alleged reason given to participants was that this procedure was necessary because the study involved food and drink samples. The alleged other participant then either arrived a moment (no provocation) or 15 minutes (provocation) after the participants themselves had arrived at the laboratory and was either visible for the participants next to the door or invisible behind the door. Participants then overheard the following dialog between the experimenter (E) and the confederate (C):

C: "Hello, I am here for the study on personality and taste preferences." E: "I'll bring you to the lab where another experimenter is waiting for you." (no provocation)

C: "Hello, I am here for the study on personality and taste preferences." E: "You are very late." C: "It's just a study. I can be late for that." E: "But the other participant had to wait for you." C: "She/he can wait a bit. She/he shouldn't make a fuss about it." E: "I'll bring you to the lab where another experimenter is waiting for you." (provocation)

After this dialog the experimenter allegedly brought the confederate to an adjacent room and thus the confederate was no longer visible in any of the conditions. Subsequently the procedure was similar to McGregor et al.'s (1998) second study and the procedure used in Study 1. However, in the present study participants filled out the 10-item State Self-Control Capacity Scale (Twenge et al., 2004) after filling out the taste preferences questionnaire that was handed to them by the experimenter when he/she returned from ostensibly escorting the alleged other participant to an adjacent room. Also, the instructions were adopted to make them better suited for the choice based paradigm (full instructions see Appendix 2). Furthermore, participants were always provided with response options and they were made believe that the alleged other participant would prepare a food sample for them while they were preparing a food sample for this person. Participants were told that the samples would be exchanged at the same time after they both prepared them for each other. However, participants actually never received a food sample. After preparing a sample and handing it to the experimenter, participants answered open-ended questions concerning the reasons behind their choice, and their assumptions about the study's purpose. Then they filled out the items to measure the intention behind their juice choice, and 5-point scale control and manipulation check questions on how likable the other participant was with the poles *very likable* (1) and *not likable at all* (5), their memory for what the other participant looked like with the poles *rather good memory*

(1) and *rather bad memory* (5), how provoked they felt by the behavior of the other participant with the poles *not provoked* (1) and *extremely provoked* (5) as well as if the other participant was on time for the study with the poles *rather punctual* (1) and *rather unpunctual* (5). Finally, participants filled out the German translation of the Buss-Perry Aggression Questionnaire (Herzberg, 2003) and provided demographic information (age, sex, profession, major if they indicated student as profession, and first language). Then participants were thoroughly debriefed as described in Study 1.

6.2 Results.

Manipulation check

Participants in the provocation conditions felt more provoked, $t(68)^{14} = -5.34$, $p < .01$, liked the confederate less, $t(68) = -5.99$, $p < .01$, and perceived the confederate as being more unpunctual, $t(69) = -7.69$, $p < .01$, than participants in the no provocation conditions suggesting an effective manipulation. For visible victims, participants indicated a better memory for what the confederate looked like than for non-visible victims suggesting an effective manipulation of the visibility factor as well, $t(68) = 8.11$, $p < .01$.

Juice choices

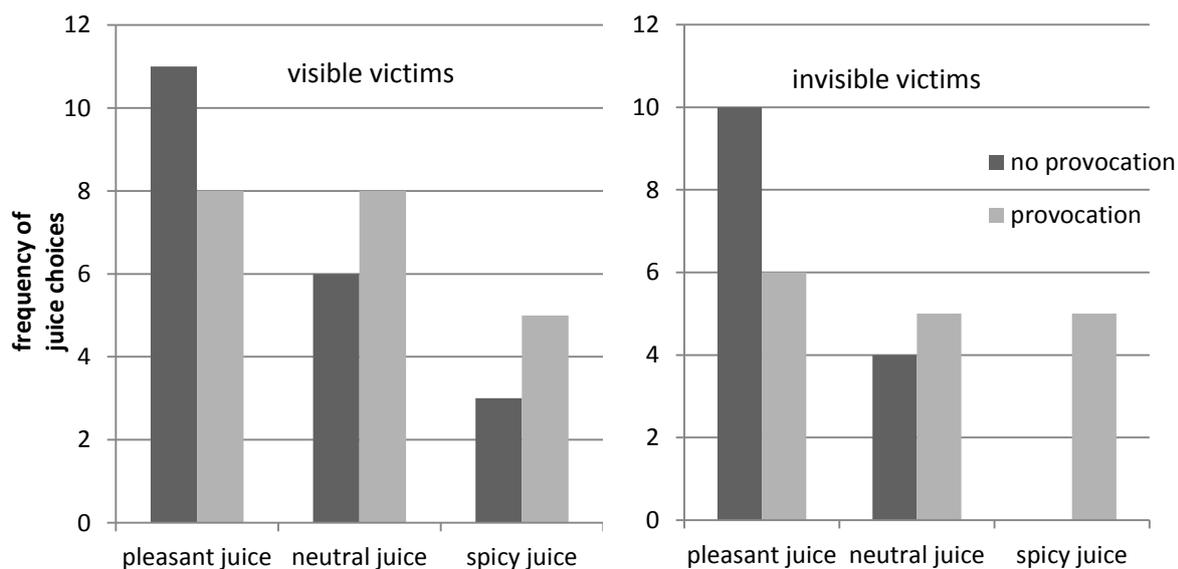
The main dependent variable in Study 4 was the juice choice. In regards to the relative frequency male participants chose the spicy juice and the neutral more often and the pleasant juice less often than female participants (spicy juice choices_{male} = 26.32 percent vs. spicy juice choices_{female} = 15.38 percent, neutral juice choices_{male} = 26.32 percent vs. neutral juice choices_{female} = 30.77 percent, pleasant juice choices_{male} = 36.84 percent vs. pleasant juice choices_{female} = 53.85). However, this difference in

¹⁴ For the manipulation check question on subjective provocation, liking of the confederate and memory for the confederate's looks 1 participant per question failed to provide an answer resulting in 1 missing value per question, $n = 70$

juice choices was not significant, $\chi^2 (2) = 1.88, p > .10$. Thus, in all further analyses data for both genders was collapsed and analyzed together.

To analyze the effect of provocation and victim visibility on choices I conducted multiple χ^2 -tests and multinomial logistic regressions. Collapsed across both visibility conditions, there was a marginally significant difference in the distribution of the juice choices between provoked and non-provoked participants, $\chi^2 (2) = 5.44, p = .07$, with spicy and neutral juice choices occurring more often in the provocation than in the no provocation conditions. I found no significant difference in the distribution of the juice choices for visible versus non-visible victims collapsed across both provocation conditions, $\chi^2 (2) = .34, p > .10$. Analyzed separately for the visible and non-visible victims, there was a significant different distribution of the juice choices between provoked versus non-provoked participants for non-visible victims, $\chi^2 (2) = 6.00, p < .05$, but not for visible victims, $\chi^2 (2) = 1.24, p > .10$ (see

Figure 3



Frequency of pleasant, neutral and spicy juice choices for provoked and non-provoked participants with visible and non-visible victims

Figure 3). For visible victims spicy juice choices occurred in both the provocation and no-provocation conditions. For non-visible victims spicy juice choices only occurred in the provocation condition.

In addition, I conducted a multinomial logistic regression analyses on the probability of juice choices with the pleasant juice as reference category for the analyses. As predictors in the model, I included the provocation (provocation vs. no provocation) as well as the visibility factor (not visible vs. visible victim). However, I did not include the interaction term between these two variables to avoid instable estimates due to a quasi-complete separation of data points (e.g., zero spicy juice choices in the victim not visible, no provocation group). Overall, the analyses revealed a marginally significant effect of provocation on the juice choices, Wald $\chi^2(2) = 5.14, p = .08$. The probability of a spicy compared to a pleasant juice choice was significantly higher with provocation in comparison to without provocation, Wald $\chi^2(1) = 4.74, p < .05$.

Mediation analysis

I hypothesized that the direct effect of provocation on spicy juice choices is mediated by participants' harmful intentions. To test this prediction I conducted an additional regression and multinomial logistic regression analysis following Iacobucci's (2012) recommendation for mediation analyses with categorical variables. I estimated a regression model on harmful intentions and included the provocation (provocation vs. no provocation) and visibility factor (not visible vs. visible victim) as predictors. The results indicated a significant connection between the possible mediator harmful intention and the predictor provocation, $b = .58, t(1) = 3.22, p < .01$. Provoked participants had stronger harmful intentions than non-provoked ones. The visibility of the victim was not significantly related to participants' harmful

intentions. The full multinomial logistic regression model on juice choices as criteria (reference category pleasant juice), with the predictors provocation (provocation vs. no provocation) and visibility (not visible vs. visible victim) as well as the mediator harmful intentions, revealed a significant overall effect of harmful intentions on juice choices, Wald $\chi^2(2) = 14.57, p < .01$. The overall effect of provocation, however, was no longer significant, Wald $\chi^2(2) = 1.01, p > .10$. The probability of choosing the spicy in comparison to the pleasant juice was significantly higher the more harmful the participants' intentions were, Wald $\chi^2(1) = 14.38, p < .01$. All other effects were not significant, $p > .10$ (see Table 3). Calculating the z -value for the mediation effect, as suggested by Iacobucci (2012), revealed a significant mediation of the provocation effect by the participants harmful intentions, $z_{\text{mediation}} = 2.41, p < .001$.

Table 3

Coefficients and standard errors for multinomial logistic regressions for predicting the juice choice (reference category: pleasant juice) with predictors provocation (no provocation vs. provocation) and visibility (not visible vs. visible victim) as well as harmful intention and for the OLS regression predicting harmful intentions with the predictors provocation (no provocation vs. provocation) and visibility (not visible vs. visible victim).

multinomial logistic regression: direct path				regression model (intention as criteria)			multinomial logistic regression: full model			
	juice type	estimate	SE		estimate	SE	juice type	estimate	SE	
intercept	neutral	-.92	.51	intercept	1.86**	.17	intercept	neutral	-1.77	2.12
intercept	spicy	-2.16**	.76	provocation	.58**	.18	intercept	spicy	-9.03	3.07
provocation	neutral	.68	.54	visibility	.14	.18	provocation	neutral	.53	.56
provocation	spicy	1.63*	.75				provocation	spicy	-.10	1.08
visibility	neutral	.30	.55				visibility	neutral	.18	.56
visibility	spicy	.36	.69				visibility	spicy	.25	1.01
							intention	neutral	.49	.49
							intention	spicy	3.06**	.81

Note. * regression coefficient significant at $p < .05$

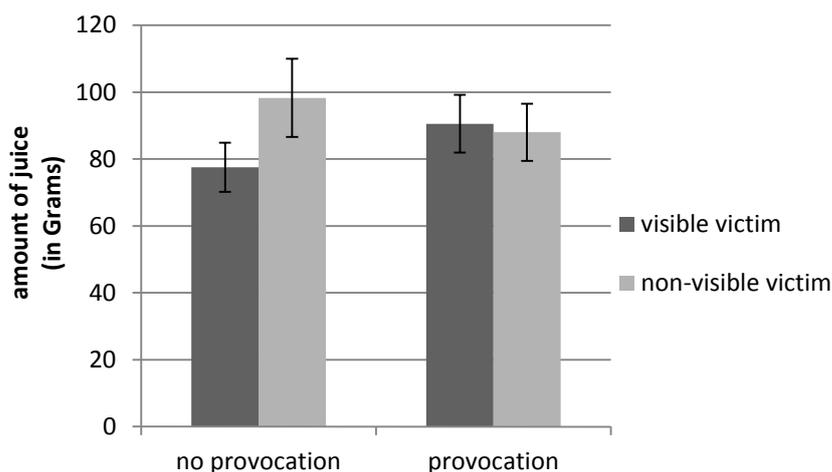
** regression coefficient significant at $p < .01$

Allocated amount

On average male participants allocated more juice, $M = 99.32$, $SD = 41.45$, than female participants, $M = 83.65$, $SD = 35.31$. However, this difference was not significant, $t(69) = 1.58$, $p > .10$. Thus, in all further analyses data for both genders was collapsed and analyzed together.

To examine the allocated amount of juice I conducted a general linear model with the allocated juice as criteria and the predictors type of chosen juice (pleasant vs. neutral vs. spicy), visibility (not visible vs. visible victim) and provocation (provocation vs. no provocation) as well as all pair wise and triple interactions. The analyses revealed an interaction between provocation and visibility of the victim, $F(1, 60) = 4.18$, $p < .05$ (see Figure 4). However, a simple effect analysis controlling for multiple tests neither revealed a significant simple effect of visibility nor provocation, all $p > .10$.

Figure 4



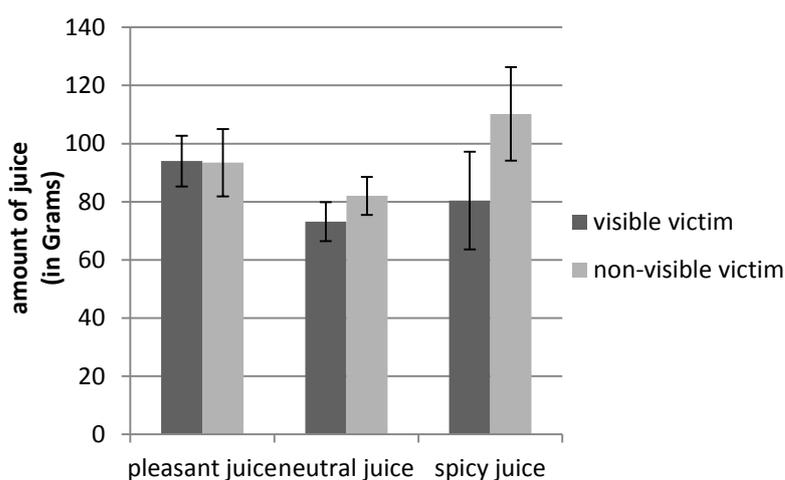
Average allocated amount of juice for visible and non-visible victims with provoked and non-provoked participants, SE as error bars

In addition, the analysis revealed a marginally significant interaction between type of chosen juice and the visibility factor, $F(2, 60) = 2.59, p = .08$ (see Figure 5). However, a simple effect analysis controlling for multiple tests neither revealed a significant simple effect of type of juice nor visibility, all $p > .10$. All other effects were non-significant ($p > .10$).

Self-control

In Study 2 and 3 I failed to replicate effects of provocation on state self-control demonstrated by Denson et al. (2011) with a rather low intense form of provocation. I conducted a t -test to assess whether I could find a difference in state self-control measured with the 10-item State Self-Control Capacity Scale (Twenge, Muraven, & Tice, 2004) between provoked and non-provoked participants right after the interpersonal provocation used in Study 4. However, again there was no significant difference in self-control between provoked and non-provoked participants $t(67) < 1, p > .10$.

Figure 5



Average allocated amount of the chosen pleasant, neutral or spicy juice with visible and non-visible, SE as error bars

Trait aggression

In two ANOVAs with the overall trait aggression score or the physical trait aggression score of the German version of the Buss-Perry Aggression questionnaire (Herzberg, 2003) as dependent variables respectively and the chosen juice (pleasant vs. neutral vs. spicy) as factor there was no significant difference, neither in overall trait, $F(2, 66) < 1, p > .10, n = 68$ due to missing values, nor in physical aggression, $F(2, 68) < 1, p > .10, n = 70$ due to a missing value, for participants choosing the pleasant, neutral or spicy juice.

6.3 Discussion

With the interpersonal provocation used in Study 4 the amount of spicy juice choices was successfully increased. Out of 71 participants, 13 chose to allocate the spicy juice and thus the modified paradigm is applicable in aggression research.

In Study 4, I demonstrated that the type of chosen juice was related to the provocation factor. Provoked participants chose the spicy juice more often than non-provoked ones. Moreover, the results indicated that the effect of provocation was mediated by participants' harmful intention. This result links the juice choices to a key definitional aspect of aggression and thus speaks in favor of the validity of juice choices as an aggression measure when participants are provided with response options (Tedeschi & Felson, 1995; Ritter & Eslea, 2005).

I did not find the predicted lower choice rate of the spicy juice with visible in comparison to not visible victims. Judging on basis of the descriptive pattern, the opposite was the case but this difference in the choice distributions was not significant. The descriptive difference that is not in line with previous research might be caused by a different kind of operationalization compared to previous research

since the participants did not see the victim while performing the aggressive behavior in both conditions.

With the fourth study I also replicated the finding of Study 1 that the presence or absence of a provocation had no effect on the amount of juice applied with choice options. This might reflect lowered demand characteristics of the modified paradigm in comparison to the traditional Hot Sauce Paradigm.

Furthermore, in Study 4 I replicated the null findings of Study 2 and 3 in regard to an effect of provocation on the participants' self-control. With a different, more intense interpersonal provocation I was not able to replicate previous research indicating a higher depletion of provoked compared to unprovoked participants (Denson et al., 2011). This failure to replicate previous findings in three studies using two different forms of provocation methods questions the generalizability of the effect documented in previous research.

7 Study 5

In Study 4 I documented a connection between harmful intentions and juice choices in the modified Hot Sauce Paradigm. The mediation of the provocation effect on spicy juice choices by participants' harmful intentions, demonstrated in Study 4, indicated the validity of the juice choices as an aggression measure. To examine the validity of the modified Hot Sauce Paradigm further and to compare it with the original version of the paradigm, I analyzed the intentions behind participants' juice choices in Study 5. An action has to be perceived as intentional harm-doing to be perceived as aggression (Dollard et al., 1939; Berkowitz, 1962; Baron & Richardson, 1994). Students, for example, rated actions described in different scenarios as more aggressive when the aggressor had harmful in comparison to instrumental or altruistic intentions (Berkowitz, Mueller, Schnell, & Padberg, 1986). With Study 5 I therefore

wanted to focus on the perception of spicy juice choices from the victims' point of view. Ethical constraints, however, prevented me from actually forcing participants to drink a large amount of the very unpleasant, spicy juice. Thus, I conducted a scenario study which was based on the procedure of Study 4. Participants were asked to imagine being in the role of the target that received a spicy juice sample from another participant. I manipulated whether or not the person that allocated the juice had choice options. Hence, I contrasted the original Hot Sauce Paradigm with the modified choice option version. This allowed me to test whether the inclusion of choice options heightens the paradigm's validity by strengthening the perceived link between harmful intentions and the juice allocation. With harmful intention being a key element for the classification of an action as aggression, a strengthened connection between the observed behavior in the modified Hot Sauce Paradigm and harmful intention would suggest an improved validity for measuring aggression.

I hypothesized that the juice allocation would more likely be perceived as based on harmful intentions in the modified choice-based paradigm than in the traditional Hot Sauce Paradigm.

7.1 Method

Participants and design

Participants were 60 students and people approached in front of the psychology department (14 males). The design of the study was a three experimental group design with the between-subject groups (no choice options vs. non-explicit choice options vs. explicit choice options). I measured the perceived harmful intentions behind the juice allocation from the victim's perspective as the dependent variable in this study.

Material and procedure

The scenario imagined by participants was based on the procedure of Study 4 without the victim visibility factor. Also, I asked all participants to imagine that they arrived 15 minutes late for their scheduled appointment so that the other participant had to wait for them to arrive. Then participants were instructed to imagine that the other participant would prepare a juice sample for them. Participants were informed that while doing so the other participant saw a questionnaire indicating their taste preferences that they had just filled out. At this point in the study I presented a filled-out questionnaire, similar to the one used in all the studies before, to participants indicating that they do not like spicy food and drinks at all. Then I instructed participants to imagine that they received a cup half-full with a juice that was selected by the other participant. To manipulate the availability of choice options participants were either instructed to imagine that the other person chose the amount of the juice, chose the juice out of three juices, or chose the juice out of three juices with neutral, pleasant and spicy taste and thus the choice options were stated explicitly in the last condition. After that all participants tried a sample of the spicy juice and answered the following questions on 5-point scales to measure the perceived intentions of the other participant “What expectations do you think would the other person have in regards to your liking of the juice?”, “What do you think, how much would the other person like it, if you would not like the juice at all?” and “What do you think, how much would the other person like it, if you would really like the juice?” with the endpoints *no liking at all* (1) and *extreme liking* (5). I computed an intention measure by reverse scoring item 2, summarizing and averaging these three items (Cronbach’s $\alpha = .83$). Additionally, I asked participants how likable they thought the other person was on a 5-point scale with the endpoints *very likable* (1) and *very unlikable* (5), how much they liked the juice and how much they thought the other person would like the juice

with the latter two answered on a 5-point scale with the endpoints *no liking at all* (1) and *extreme liking* (5) as well as demographic questions (age, sex, profession, major if they indicated student as profession).

7.2 Results

I conducted an ANOVA with the factor choice options (no choice option vs. non-explicit choice options vs. explicit choice options) on the perceived intentions as dependent variable to test whether there was a difference in the perceived harmful intentions between the three experimental groups. The analysis revealed that an overall difference between the experimental groups did not reach conventional levels of significance, $F(2, 57) = 2.17, p = .12$. Since I hypothesized a difference between experimental groups with and without choice options, I conducted a planned orthogonal contrast between the no choice-option condition and the choice-option conditions (non-explicit choice options and explicit choice options). This contrast indicated a marginally significant difference between the no choice-option condition and the choice-option conditions, $F(1, 57) = 3.46, p = .07$. With choice options participants perceived the intentions underlying the juice allocation as more harmful reflected in lower means on the intention measure, $M_{\text{no choice}} = 2.63, SD = 1.26$; $M_{\text{non-explicit choice}} = 2.25, SD = 1.01$; $M_{\text{explicit choice}} = 1.93, SD = .89$. On all other variables, there were no significant differences between the groups.

7.3 Discussion

Extending the result of Study 4, which indicated a link between spicy juice choices and participants' harmful intentions, the results of Study 5 gained further support for an improved validity of the Hot Sauce Paradigm due to the inclusion of choice options. Participants perceived a spicy juice choice more as an intentional harm-doing than the choice of just the amount of a preselected spicy juice. Thus, the

inclusion of choice options not only corrects a confirmation tendency in research due to only including the behavioral option that the research focus lies upon, but also the participants' behavior in the modified Hot Sauce Paradigm is more closely related to the key aggression-defining construct of harmful intentions (Tedeschi & Felson, 1995).

8 General Discussion

In this concluding general discussion I will review and interpret the result of the presented studies as well as discuss methodological limitations, future research directions and practical implications for aggression research, and also for psychology in general.

8.1 Summary of the results

Choice options

In comparison to classical paradigms used to study aggression in the laboratory, modern laboratory paradigms, especially the Hot Sauce Paradigm, have been considerably improved to study aggression due to a revision of severe shortcomings impairing classical paradigms. The aggression measure gained with the Hot Sauce Paradigm is not biased due to a prosocial or competitive cover story and has a higher external validity because it is based on real cases in which the allocation of spicy food was used as a form of aggression (Ritter & Eslea, 2005; Lieberman et al., 1999). However, even modern laboratory research paradigms, including the Hot Sauce Paradigm, often suffer from an unavailability of non-aggressive response alternatives for participants. Without choice options it is not possible to test under which conditions someone will show aggressive behavior but only how much aggression someone would show if he/she decides to act aggressively in certain situation. Further shortcomings of laboratory aggression research are an unvaried large distance between the aggressor and the (alleged) victim. Additionally, limited

attention has been paid to key mediating variables like participants' intentions, affect or cognitions (Ritter & Eslea, 2005). With 5 studies presented in this dissertation, I attempted to answer the overriding question whether or not the Hot Sauce Paradigm (McGregor et al. 1998; Lieberman et al., 1999), as one commonly used modern laboratory research paradigm, could be improved through modifying aspects related to these shortcomings. In general, over the course of five studies I not only found evidence questioning the traditional paradigm's validity but also evidence for an improvement in the validity of the paradigm when choice options were included.

Based on previous research (Barlett et al., 2009; Böhm & Streicher, 2010), I tried to analyze the effect of missing response alternatives on the behavior shown in the Hot Sauce Paradigm in Study 1. By including conditions with choice options and contrasting them with conditions in which participants were only provided with either a single pleasant, neutral or aggressive option, I isolated the effect of choice options on the behavior shown by participants. With the conditions in which participants were provided with response alternatives, I tested whether this shortcoming, potentially leading to an underestimation of different, possible prosocial behavior as well as biased aggression estimates, could be improved. In Study 1 I replicated the classical Hot Sauce Paradigm finding of larger amounts of spicy juice allocated by provoked than unprovoked participants, previously interpreted as more aggressive behavior. However, the results of this study also indicated that the effect of provocation on the allocated amount was not bound to just the unpleasantly spicy juice. In fact, there was a main effect of the provocation manipulation on the allocated amount of juice with provoked participants allocating more juice than non-provoked ones. This finding questions the amount of allocated spicy juice as a valid measure for aggression. Provoked participants might have a higher responsiveness to the paradigms demand in the direction of applying more sauce, which might be the reason for this data pattern.

With choice options the difference in the allocated amount between provoked and non-provoked participants was descriptively the smallest and a simple effect revealed that the difference in the allocated amount was not significant for provoked and non-provoked participants in conditions with choice options. With choice options provided to participants the demand of the paradigm is lower since the attention focus is no longer drawn to one behavioral option. Thus, this might explain that there was no significant difference in the allocated amount between provoked and non-provoked participants in conditions with choice options.

Self-control

As one explanation for a potentially increased responsiveness to the paradigm's demand of provoked participants in comparison to unprovoked ones, I proposed lowered self-control due to rumination about the provoking episode (Denson et al., 2011). I tested this possible explanation with Study 2, 3, and 4.

In Study 2, I orthogonally manipulated the demand characteristics of the instruction (regular instruction vs. less demand impaired instruction) and provocation (no provocation vs. provocation) as well as measured participants' state self-control. In Study 3, I orthogonally manipulated the easiness of behaving in line with the paradigm's demand and apply large amounts of the spicy juice (easy vs. hard application) and provocation (no provocation vs. provocation) as well as measured participants' state self-control. In Study 4, I measured participants' state self-control after a different, more intense interpersonal provocation. In Study 2, there was a significant difference in the allocated amount of juice between provoked and non-provoked participants when the data was collected with regular instructions but not when it was collected with the less demand-prone instructions. This suggests that the data pattern of Study 1 might be due to the demand created by the wording of the instruction and by not providing response alternatives. Taken together, the results of

Study 1 and 2 question the validity of the Hot Sauce Paradigm for measuring aggressive behavior.

In regard to the proposed self-control explanation for the higher responsiveness to the paradigm's demand, in both Study 2 and 3 I was not able to replicate the previously found lowered state self-control after a provoking event (Denson et al., 2011). Even with a different, more intense interpersonal provocation method used in Study 4 that included that participants waited 15 minutes for another alleged participant to arrive as well as rather rude behavior toward them after the arrival I could not find a difference in self-control between provoked and non-provoked participants after the provoking event occurred. These findings question the generalizability of previous research on the depleting effects of provocation (Denson et al., 2011). On the basis of these results it is quite unlikely that the difference between provoked and non-provoked participants' responsiveness to the paradigm's demand is due to a difference in self-control. Further research is necessary to clarify what the exact process is that caused the data pattern of Study 1 and 2. One alternative explanation might be that participants indeed ruminate about the provoking event and that this distracting rumination makes them follow the demand more so compared to non-provoked participants but that this process is not intense enough to actually deplete them. Another potential alternative explanation might be that participants felt ostracized (e.g., Williams, 1997) after the provocation manipulation and due to this feeling were more likely to follow experimental demand. Previous research indicated that ostracism not only leads to increased aggression (Twenge, Baumeister, Tice, & Stucke, 2001) but also increases conformity (Williams, Cheung, & Choi, 2000) and social information processing (Pickett, Gardner, & Knowles, 2004). Thus, after the provocation participants might have shown a higher responsiveness to the demand and applied more juice because they felt ostracized by the behavior of their alleged

experimental partner. Clearly, the provocation method used in Study 1, 2 and 3 is different from the typical experimental methods that are used to socially exclude an individual. Commonly ostracism is operationalized with a (cyber-)ball tossing game in which participants are either included (no ostracism condition) or are excluded after the first few tosses and do not receive the ball anymore (ostracism condition; Williams et al., 2000; Williams & Sommer, 1997). Even though the provocation manipulation was clearly different from such an operationalization, provoked participants in Study 1, 2 and 3 might have felt rejected when the alleged other participant in his/her first interaction with them chose a very sour juice for them out of a selection that range from not sour to very sour without any apparent reason to do so. Thus, feeling rejected and (mildly) ostracized might have increased their need to act on the demand of the paradigm and behave in a way that they thought the experimenter wanted them to behave to fulfill their need of belonging and affiliate with the only other person that was there. However, this is a post-hoc explanation that needs to be tested in additional research.

Improved paradigm

In addition to showing an impairment of the traditional Hot Sauce Paradigm due to demand and missing response options in Study 1 and 2, I was able to improve the paradigm's validity with the inclusion of choice options. The amount of aggression that I observed in Study 1 in the conditions using the modified choice-based paradigm was very low (2 of 38 participants chose the spicy juice) but the dependent measure was no longer biased since participants were no longer detracted from their default choice option prosocial behavior. In Study 4, in which I applied a more intense interpersonal provocation to heighten the amount of aggressive choices, all data was collected with the modified version of the paradigm that provides choice options to participants. In this study 13 of 71 participants chose the spicy juice and

thus I increased the number of aggressive choices making the paradigm applicable in aggression research. With the higher choice rate of the aggressive option certain analyses like prediction models were possible that are not applicable when the crucial response occurs very seldom. In Study 4, I was able to demonstrate a connection between spicy juice choices and harmful intentions. In fact, the effect of provocation on spicy juice choices was mediated by the participants' harmful intention. The chosen type of mediation analyses might be criticizable since not all possible variables that might mediate the process were tested and the mediator was not manipulated experimentally. However, harmful intention as mediator has a strong theoretical foundation since it is one of the key concepts in the definition of aggression justifying this type of analysis (e.g., Dollard et al., 1939; Berkowitz, 1962; Baron & Richardson, 1994; Tedeschi & Felson, 1994) validating the spicy juice choice as an aggression measure.

In Study 5, the connection between harmful intentions and juice allocation also was perceived as stronger by participants imagining that they would receive a juice sample from someone who had choice options (modified version) in comparison to from someone who could only vary the amount of a preselected spicy juice (classical Hot Sauce Paradigm). Thus, the inclusion of choice options strengthened the perceived link between the allocation behavior and concepts critically for a valid aggression measure.

Underestimation of non-aggressive behavior

The amount of aggression that I observed in the choice-based paradigm was very low in the first study (2 of 38 participants chose the spicy juice). Due to the interpersonal provocation that I applied in Study 4, the amount was higher (13 of 71 participants chose the spicy juice). However, the frequency of the behavior was still

rather low. These small frequencies demonstrate the large occurrence of non-aggressive behavior that is underestimated with paradigms that do not include response options and only allow participants to show a varying degree of aggressive behavior.

Providing alternative behavioral options is not just important in aggression research. Although in this field, with a focus on a specific form of behavioral outcome a biasing tendency to only provide the option the research focus lies upon seems likely (Fiedler, 2011). As mentioned in the introduction, another example of a confirmation tendency (Ferguson, 2007) in the field of aggression psychology is the research on effects of video games. The vast majority of research in this field focused on demonstrating negative effects (Barlett et al., 2009). By seldom including dependent variables that could indicate beneficial effects of videogames possible positive outcomes of violent or prosocial games have largely been ignored in past research and only recently there have been improvements (Greitemeyer & Oswald, 2012).

To avoid positive testing – a hypothesis testing strategy in which the focus of the information search lies on the property, process or event that is expected to occur while additional alternative information is neglected (Klayman & Ha, 1987; Fiedler et al., 1999) – in aggression psychology but also in all other research fields researchers should provide non-biased response options or multiple dependent variables capturing more than the hypothesized events, process and properties (Fiedler, 2011). The research documented in this dissertation is just one example of how observed data patterns can change dependent on the provided response options and, with that, the conclusions based on the data.

A possible explanation for why aggressive choices occurred infrequently might be that participants have to transgress a higher threshold for choosing an aggressive option in comparison to just administering a larger amount of a preselected aggressive option. Bearing this in mind, the behavior observed in the choice-based modified paradigm might be linked more closely to rather controlled, instrumental forms of aggressive behavior than the behavior shown in the traditional Hot Sauce Paradigm. The behavior observed in the traditional version, in contrast, might be associated more strongly with thoughtless, impulsive aggression that is less controlled by participants (Berkowitz, 2008). The strong relation between intention and juice choices and the non-significant relation between intention and the allocated amount of chosen spicy juice observed in Study 4 is preliminary evidence that points in this direction, $r = -.14, p < .10$. However, since the correlation is only based on $n = 13$ this evidence has to be regarded carefully.

Trait aggression

Over the course of 4 studies (Study 1-4) in which I measured general trait aggression and physical trait aggression either with a German 12-item short version of the Buss-Perry Aggression questionnaire (Bryant & Smith, 2001) or the regular 27-item German version of the Aggression Questionnaire (Herzberg, 2003) I did not find a connection between the allocated amount of spicy juice (Study 1, 2, 3) or the juice choice (Study 4) and trait aggression.

The evidence for a connection between the allocated amount of hot sauce and physical aggression documented in previous research (Liebermann et al., 1999) was limited to a rational thinking mode characterized by a more deliberate thinking style. No correlation was found by these authors for an experiential mode characterized by a quick, intuitive driven thinking style. In my studies, in which I measured the amount

of allocated juice, I did not control for the moderator relational versus experiential thinking mode. This might explain why I did not find a correlation. I also did not find a connection between trait aggression and juice choices. Participants who chose a neutral, pleasant or spicy juice did not differ in trait aggression. This null finding might be due to the small number of participants who chose the spicy juice ($n = 13$) especially if the effect size for the trait differences is small. However, the paradigm might also capture situational specific aggressive behavior that is – in the case of the juice choices – more strongly connected to harmful intentions induced by the situation than to trait differences. This reasoning aligns with the finding that the revenge likelihood is predicted better by contextual characteristics than personality variables (Gollwitzer, 2007). In addition, the predictive power for personality variables measured with a self-report might be small since these measures are biased in the direction of socially desirable behavior (Krahé, 2001). Nevertheless, future research should further analyze the relation between the allocated amount as well as the juice choices in the regular and modified Hot Sauce Paradigm and trait aggression variables as well as other concepts validating that the behavior shown in the paradigm reflects aggression.

8.2 Limitations and future research proposals

To further establish the validity of the choice-based paradigm many open questions, like the above mentioned connection to trait variables, still remain and have to be addressed in future research. In addition, limitations of the present research have to be resolved in future studies as well.

In Study 1, I analyzed changes in the affective state in the course of the Hot Sauce Paradigm and found differences in participants' hostility level. Possibly due to a mild form of provocation, the average hostility level was very low and thus it would be worth analyzing participants' affective state during the traditional and modified

Hot Sauce Paradigm with a different, more intense provocation. Even though I conducted Study 4 with a more intense interpersonal provocation, I was not able to analyze affective states during the choice-based paradigm. Due to the provocation method applied in Study 4 both a baseline measure before and a measure of the affective state right after the alleged other participant arrived was not possible without making participants suspicious about the cover story of the study. Thus, the mediating role of affective states still needs to be clarified in future research.

Another problem due to provocation method in Study 4 was that I had to exclude many participants because they indicated doubts that an actual participant would behave this rudely after being late for the study and, possibly due to the unusual situation, considered the possibility that the study's real purpose was to analyze revenge behavior. Thus, the modified Hot Sauce Paradigm should be combined with alternative provocation methods that do not suffer from these limitations in further research.

In Study 2 I found evidence for an influence of cues in the instruction on the behavior shown in the Hot Sauce Paradigm. A difference in the allocated amount of spicy juice between provoked and unprovoked participants only emerged with a regular instruction and not with one that did not explicitly address the amount of the juice that the participants were instructed to allocate in the course of the study. To further investigate if the original Hot Sauce Paradigm assesses aggression in the sense of intentionally harming another person or if the behavior is mostly affected by cues in the instruction, in a future study participants could first receive a full glass of spicy juice and then they could be instructed to let a certain amount of this juice out of the glass. The instruction could be formulated with a focus on letting a certain amount of the juice out of the cup that the other participant would allegedly receive. Thus, the goal could be to create an instruction that makes letting a rather large amount of juice

out of the cup the salient behavioral option for participants. A participant that would be compliant and follow the demand would allocate a lower amount of the spicy juice to the other alleged participant and hence would be less aggressive. Based on the results of Study 1 this might especially be the case for provoked participants and could be tested in future research.

Especially interesting future research questions might be questions that would lead to different predictions for the regular and the modified Hot Sauce Paradigm. Recently, Gollwitzer and Denzler (2009) demonstrated that the motives behind revenge behavior are more about making the target understand that his or her behavior was morally wrong instead of making another person suffer. Participants experienced more goal fulfillment when the target signaled understanding after the revenge. Applying the regular versus choice-based paradigm to answer research questions arising from this finding would lead to different predictions. Based on the results of Study 4 and 5 it can be predicted that participants would assume that it is more likely to attribute aggressive or prosocial meaning to the juice allocation when they are provided with the possibility to choose a juice and are not just able to choose the amount of a preselected juice.

9 Conclusion

The research of my dissertation highlights the importance of providing non-aggressive response options to participants in laboratory aggression research. The validity of the Hot Sauce Paradigm, a commonly used laboratory paradigm, was improved due to the inclusion of choice options. With such options, aggressive choices were connected to a key aggression defining element: participants' harmful intentions thus making the juice choices a valid measure for behavioral aggression. Particularly if the ultimate goal is to infer conditions under which a certain behavior occurs participants have to be able to make a choice of what kind of behavior they

want to show. My modified Hot Sauce Paradigm constitutes a research instrument to achieve this goal.

References

- Adachi, P. J. C., & Willoughby, T. (2011). The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence? *Psychology of Violence, 1*, 259-274.
- Ahmed, S. M. (1979). Visibility of the victim. *The Journal of Social Psychology, 107*, 253-255.
- Anderson, C. A., Benjamin, A. J. Jr., & Bartholow, B. D. (1998). Does the gun pull the trigger? Automatic priming effects of weapon pictures and weapon names. *Psychological Science, 9*, 308-314.
- Anderson, C. A., & Bushman, B. J. (1997). External validity of “trivial” experiments: The case of laboratory aggression. *Review of General Psychology, 1*, 19-41.
- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51.
- Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science, 8*, 3-9.
- Atkinson, C., & Polivy, J. (1976). Effects of delay, attack, and retaliation on state depression and hostility. *Journal of Abnormal Psychology, 85*, 570 – 576.
- Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.

- Bandura, A. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality*, 9, 253-269.
- Barlett, C. P., Anderson, C. A., & Swing, E. L. (2009). Video game effects – confirmed, suspected, and speculative. A review of the evidence. *Simulation & Gaming*, 40, 377-403.
- Barlett, C. P., Branch, O., Rodeheffer, C., & Harris, R. (2009). How long do the short-term violent video game effects last? *Aggressive Behavior*, 35, 225-236.
- Baron, R. A. (1979). Heightened sexual arousal and physical aggression: An extension to females. *Journal of Research in Personality*, 13, 91-102.
- Baron, R. A., & Richardson, D. R. (1994). *Human aggression*. (2nd ed.). New York: Plenum Press.
- Berkowitz, L. (1962). *Aggression: A social psychological analysis*. New York: McGraw-Hill.
- Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. New York, NY, England: McGraw-Hill Book Company.
- Berkowitz, L. (2008). On the consideration of automatic as well as controlled psychological processes in aggression. *Aggressive Behavior*, 34, 117-129.
- Berkowitz, L., Cochran, S. T., Embree, M. C. (1981). Physical pain and the goal of aversively stimulated aggression. *Journal of Personality and Social Psychology*, 40, 687-700.

- Berkowitz, L., Corwin, R., & Heironimus, M. (1962). Film violence and subsequent aggressive tendencies. *Public Opinion Quarterly*, *27*, 217-229.
- Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, *7*, 202-207.
- Bettencourt, A. B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin*, *119*, 422-447.
- Berkowitz, M. W., Mueller, C. W., Schnell, S. V., & Padberg, M. T. (1986). Moral reasoning and judgments of aggression. *Journal of Personality and Social Psychology*, *51*, 885-891.
- Bluemke, M., & Zumbach, J. (2007). Implicit and explicit measures for analyzing the aggression of computer gamers. In G. Steffgen & M. Gollwitzer (Eds.) *Emotions and aggressive behavior* (pp. 38-57). Göttingen, Germany: Hogrefe & Huber.
- Böhm, F., & Streicher, B. (2010). Fette Aggressionsmessung: Eine geschlechtssensitive Erweiterung des Hot-Sauce-Paradigmas. Poster presented at the *47th biannual meeting of the German Psychology Society (DGPs)*, University of Bremen, Germany.
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research in Personality*, *35*, 138-167.

- Bushman, B. J. (1995). Moderating role of trait aggressiveness in the effects of violent media aggression. *Journal of Personality and Social Psychology*, *69*, 950-960.
- Buss, A. H. (1961). *The psychology of aggression*. New York: Wiley.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*, 452-459.
- Caprara, G.V., Renzi, P., Alcini, P., D'Imperio, G., & Travaglia, G. (1983). Instigation to aggress and escalation of aggression examined from a personological perspective: The role of irritability and of emotional susceptibility. *Aggressive Behavior*, *9*, 345-351.
- Carlson, M., & Miller, N. (1988). The differential effects of social and nonsocial negative events on aggressiveness. *Sociology and Social Research*, *72*, 155-158.
- Carlson, M., Marcus-Newhall, A., & Miller, N. (1989). Evidence for a general construct of aggression. *Personality and Social Psychological Bulletin*, *15*, 377-389.
- Cherek, D. R. (1981). Effects of smoking different doses of nicotine on human aggressive behavior. *Psychopharmacology*, *75*, 339-345.
- Denson, T. F., Pedersen, W. C., Friese, M., Hahm, A., & Roberts, L. (2011). Understanding impulsive aggression: Angry rumination and reduced self-control capacity are mechanisms underlying the provocation-aggression relationship. *Personality and Social Psychology Bulletin*, *37*, 850-862.

- Denzler, M., Förster, J., & Liberman, N. (2009). How goal-fulfillment decreases aggression. *Journal of Experimental Social Psychology, 45*, 90-100.
- DeWall, C. N., Baumeister, R. F., Stillman, T. F., Gailliot, M. T. (2007). Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental Social Psychology, 43*, 62-76.
- Diener, E., Fraser, S. Beaman, A., & Kelem, R. (1976). Effects of deindividuation variables on stealing among Halloween trick-or-treaters. *Journal of Personality and Social Psychology, 33*, 178-183.
- Dollard, J., Doob, L. W., Miller, N. E., Mower, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.
- Evers, C., Fischer, A. H., Mosquera, P. M., & Manstead, A. S. (2005). Anger and social appraisal: A “spicy” sex difference? *Emotion, 5*, 258-266.
- Fahrenberg, J. (2006). *Self-reported state – single items or scales like AD-ACL or PANAS? Ambulatory Assessment – European Network. Statements and Open Peer Commentaries*. Retrieved from http://www.ambulatory-assessment.org/typo3/ambulatory/fileadmin/user_upload/statements/Self_report_of_Subjective_State_091106.pdf
- Ferguson, C. (2007). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior, 12*, 470-482.
- Fiedler, K. (2011). Voodoo correlations are everywhere – not only in neuroscience. *Perspectives on Psychological Science, 6*, 163-171.

- Fiedler, K., Walther, E., & Nickel, S. (1999). The auto-verification of social hypotheses: Stereotyping and the power of sample size. *Journal of Personality and Social Psychology, 77*, 5-18.
- Geen, R. G. (2001). *Human Aggression*. (2nd edition). Buckingham: Open University Press.
- Geniole, S. N., Carré, J. M., & McCormick, C. M. (2011). State, not trait, neuroendocrine function predicts costly reactive aggression in men after social exclusion and inclusion. *Biological Psychology, 87*, 137-145.
- Giancola, P. R., & Chermack, S. T. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior, 3*, 237-253.
- Graham, K., Tremblay, P. F., Wells, S., Parnanen, K., Purcell, J., & Jelley, J. (2006). Harm and intent and the nature of aggressive behavior: Measuring naturally-occurring aggression in barroom settings. *Assessment, 13*, 280-296.
- Greitemeyer, T., & Osswald, S. (2010). Effects of prosocial video games on prosocial behavior. *Journal of Personality and Social Psychology, 98*, 211-221.
- Gollwitzer, M. (2007). How affective is revenge? Emotional processes involved in vengeful reactions to experienced injustice. In G. Steffgen & M. Gollwitzer (Eds.), *Emotions and aggressive behavior* (pp. 115-129). Göttingen, Germany: Hogrefe & Huber.
- Gollwitzer, M., & Denzler, M. (2009). What makes revenge sweet: Seeing the offender suffer or delivering a message? *Journal of Experimental Social Psychology, 45*, 840-844.

- Gottfredson, M., & Hirschi, T. (1993). A control theory interpretation of psychological research on aggression. In R. B. Felson & J. T. Tedeschi (Eds.), *Aggression and violence: Social interactionist perspectives* (pp. 47-68). Washington, DC: American Psychological Association.
- Herzberg, P. Y. (2003). Faktorstruktur, Gütekriterien und Konstruktvalidität der deutschen Übersetzung des Aggressions-Fragebogens von Buss und Perry. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 311-323.
- Hirsh, J. B., Galinsky, A. D., & Zhong, C.-B. (2011). Drunk, powerful, and in the dark: How general processes of disinhibition produce both prosocial and antisocial behavior. *Perspectives on Psychological Science*, 6, 415-427.
- Hokanson, J. E. (1970). Psychophysiological evaluation of the catharsis hypothesis. In E. I. Megargee & J. E. Hokanson (Eds.), *The dynamics of aggression* (pp. 74-86). New York: Harper & Row.
- Holmes, D. S. (1972). Aggression, displacement, and guilt. *Journal of Personality and Social Psychology*, 21, 296-301.
- Huffington Post. (2012). *Lillian Gomez, teacher, accused of using hot sauce to punish special needs students*. Retrieved September, 2012 from http://www.huffingtonpost.com/2012/02/01/teacher-hot-sauce-punish-students_n_1246961.html
- Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, 22, 582-594.

- Ito, T., A., Miller, N., & Pollock, V. E. (1996). Alcohol and aggression: A meta-analysis on the moderating effects of inhibitory cues, triggering events, and self-focused attention. *Psychological Bulletin*, *120*, 60-82.
- Johnson, R. D., & Downing, L. L. (1979). Deindividuation and valence of cues: Effects on prosocial and antisocial behavior. *Journal of Personality and Social Psychology*, *37*, 1532-1538.
- Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211-228.
- Krahé, B. (2001). *The social psychology of aggression*. Hove: Psychology Press.
- Lee, M. (2009). UNL Sigma Chi members in court in alleged hazing case. *Lincoln Journal Star*. Retrieved September, 2012 from http://journalstar.com/news/local/unl-sigma-chi-members-in-court-in-allegedhazing-case/article_a6a305f3-2c43-5877-a1ec-59d3a445766f.html
- Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, *25*, 331-348.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. London: Tavistock.
- Mussweiler, T., & Förster, J. (2000). The sex → aggression link: A perception-behavior dissociation. *Journal of Personality and Social Psychology*, *79*, 507-520.
- McGregor, H. A., Lieberman, J. D., Greenberg, J., Solomon, S., Arndt, J., Simon, L., & Pyszczynski, T. (1998). Terror management and aggression: Evidence that

mortality salience motivates aggression against worldview-threatening others.

Journal of Personality and Social Psychology, 74, 590-605.

Norlander, T., Nordmarker, A., & Archer, T. (1998). Effects of alcohol and frustration on experimental graffiti. *Scandinavian Journal of Psychology*, 39, 201-207.

Page, R. A., & Moss, M. K. (1976). Environmental influences on aggression: The effect of darkness and proximity of victim. *Journal of Applied Social Psychology*, 6, 126-133.

Regan, K. (2012). Man charged with assaulting 6-year-old, forcing hot sauce. *Marlborough Patch*. Retrieved September, 2012 from <http://marlborough.patch.com/articles/man-charged-with-abusing-6-year-old-with-hot-sauce>

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and Social Psychology Bulletin*, 30, 1095-1107.

Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, 123, 238-259.

Ritter, D., & Eslea, M. (2005). Hot sauce, toy guns, and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior*, 31, 407-419.

Russell, G. W., Arms, R. L., Loof, S. D., & Dwyer, R. S. (1996). Men's aggression toward women in a bungled procedure paradigm. *Journal of Social Behavior and Personality*, 11, 729-738.

Stogsdill, S. (2012). Day-care workers charged with child abuse by hot sauce. *Tulsa World*. Retrieved September, 2012 from

http://www.tulsaworld.com/news/article.aspx?subjectid=12&articleid=20120416_12_0_OKEMAH220898.

- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. (5th international edition). Boston, MA: Pearson Allyn and Bacon.
- Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality*, *35*, 297-310.
- Taylor, S., & Leonhard, K. (1983). Alcohol and human physical aggression. In R. G. Geen, & E. I. Donnerstein (Eds.), *Aggression: Theoretical and Empirical Reviews* (Vol. 1, pp. 77-101). New York: Academic Press.
- Tedeschi, J. T., & Felson, R. B. (1994). *Violence, aggression, and coercive actions*. Washington, DC, US: American Psychological Association.
- Tedeschi, J. T., & Quigley, B. M. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior*, *1*, 163-177.
- Tedeschi, J. T., & Quigley, B. M. (2000). A further comment on the construct validity of laboratory aggression paradigms: A response to Giancola and Chermack. *Aggression and Violent Behavior*, *5*, 127-136.
- Tuner, C. W., Layton, J. F., & Simons, L. S. (1975). Naturalistic studies of aggressive behavior: Aggressive stimuli, victim visibility, and horn honking. *Journal of Personality and Social Psychology*, *31*, 1098-1107.
- Twenge, J. M., Baumeister, R. T., Tice, D. M., & Stucke, T. S. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology*, *81*, 1058-1069.

-
- Twenge J. M., Muraven, M., & Tice, D. M. (2004). *Measuring state self-control: Reliability, validity, and correlations with physical and psychological stress*. Unpublished manuscript, San Diego State University.
- Warburton, W. A., Williams, K. D., & Cairns, D. R. (2006). When ostracism leads to aggression: The moderating effects of control deprivation. *Journal of Experimental Social Psychology, 42*, 213-220.
- Watson, D., & Clark, L. A. (1994). *THE PANAS-X. Manual for the positive and negative affect schedule – expanded form*. Iowa City: University of Iowa.
- Williams, K. D. (1997). Social ostracism. In R. Kowalski (Ed.), *Aversive interpersonal behaviors* (pp. 133-170). New York: Plenum.
- Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology, 79*, 748-762.
- Williams, K. D., & Sommer, K. L. (1997). Social ostracism by coworkers: Does rejection lead to social loafing or compensation? *Personality and Social Psychology Bulletin, 23*, 693-706.
- Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order vs. deindividuation, impulse and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska symposium on motivation* (pp. 237-307). Lincoln: University of Nebraska Press.

Appendix 1: Regular and less demand impaired instructions in German as used in Study 2

First Instruction

Sie werden im Folgenden an einer Untersuchung teilnehmen, die den Zusammenhang zwischen Persönlichkeit und Geschmackspräferenzen untersucht. Ihre Antworten werden völlig anonym behandelt und Sie werden für die Dauer der Untersuchung mittels einer dreistelligen Nummer identifiziert, die auf all ihrem Material erscheinen wird.

Im Rahmen der Untersuchung werden verschiedene Typen von Nahrungsmitteln untersucht. Heute werden verschieden angenehme, neutrale, saure und scharfe Getränke untersucht.

Da der Leiter des Versuchs die genauen Details der Nahrungsmittelproben und die Menge dieser Nahrungsmittel, die von den einzelnen Personen probiert werden, nicht wissen darf, ist es notwendig, dass die Studienteilnehmer diese Kostproben für einander auswählen.

Obwohl wir an diesem Tag saure, scharfe, neutrale und angenehme Getränke untersuchen, sind wir auch an anderen Geschmäckern und Konsistenzen interessiert. Daher möchte ich Sie bitten, einen Fragebogen auszufüllen, in dem es um Ihre Geschmackspräferenzen geht.

Second instruction given to the participants when they receive the taste sample allergy prepared from the other participant

Der Becher enthält einen Saft der von einer anderen Versuchsperson im Nebenraum ausgewählt wurde. Diese Person konnte den Saft aus Säften mit einer Reihe von Geschmäckern auswählen, die von neutralem bis sehr saurem Geschmack reichten. Die andere Person konnte Ihnen also jeden Geschmack aus einer bestimmten Auswahl geben und dies ist der Geschmack, den sie gewählt hat. Bitte trinken Sie den Saft vollständig. Nachdem Sie das Getränk getrunken haben, bewerten Sie es bitte mittels der folgenden Fragen und füllen dann die zusätzlichen Fragen aus.

Regular instruction	Less demand impaired instruction
Nun ist es Ihre Aufgabe eine Kostprobe eines Getränks für die andere Person vorzubereiten. Sie wählen die Menge des Getränks für die Person aus, die für Sie die gerade getrunkene Saftprobe ausgewählt hat.	Nun ist es Ihre Aufgabe eine Kostprobe eines Getränks für die andere Person vorzubereiten, die für Sie die gerade getrunkene Saftprobe ausgewählt hat.
Weil Menschen oft neugierig sind bezüglich der Geschmackspräferenzen der anderen Versuchsteilnehmer, sehen Sie hier den Fragebogen zu den Geschmackspräferenzen der Person, für die Sie im Folgenden die Menge eines Getränkes auswählen sollen.	Weil Menschen oft neugierig sind bezüglich der Geschmackspräferenzen der anderen Versuchsteilnehmer, sehen Sie hier den Fragebogen zu den Geschmackspräferenzen dieser Person.

Alle Mengen von Getränken sind nützlich und daher sollen Sie so viel oder so wenig des Getränks bereit stellen wie Sie möchten.

Die Person, für die Sie die Getränkeprobe vorbereiten, muss die gesamte Menge dieser Probe trinken.

Bitte probieren Sie das Getränk mit den dafür bereitgestellten Becher bevor Sie die Getränkeprobe für den anderen Versuchsteilnehmer wählen.

Zusätzlich haben Sie auch ein Glas Wasser erhalten, so dass Sie mögliche Unannehmlichkeiten, die Sie eventuell durch das Probieren erleben, ausgleichen können.

Bitte verschließen Sie dann den Becher. Dies dient dazu damit ich als Versuchsleiter nicht sehen kann wie viel von dem Getränk Sie für die andere Person gewählt haben.

Bitte schreiben Sie dann die Nummer der anderen Versuchsperson auf den Becher die dann der Versuchsteilnehmer erhalten wird, der für Sie den gerade von Ihnen getrunkenen Saft ausgewählt hat.

Hier erhalten Sie eine Checkliste, die alle Schritte des Vorgehens beim Zusammenstellen der Getränkeprobe noch einmal enthält.

Die Person, für die Sie die Getränkeprobe vorbereiten, muss die gesamte Probe trinken.

Bitte probieren Sie das Getränk mit den dafür bereitgestellten Becher bevor Sie die Getränkeprobe für den anderen Versuchsteilnehmer wählen.

Zusätzlich haben Sie auch ein Glas Wasser erhalten, so dass Sie mögliche Unannehmlichkeiten, die Sie eventuell durch das Probieren erleben, ausgleichen können.

Bitte verschließen Sie dann den Becher. Dies dient dazu damit ich als Versuchsleiter die bereitgestellte Probe nicht sehen kann.

Bitte schreiben Sie dann die Nummer der anderen Versuchsperson auf den Becher die dann der Versuchsteilnehmer erhalten wird, der für Sie den gerade von Ihnen getrunkenen Saft ausgewählt hat.

Hier erhalten Sie eine Checkliste, die alle Schritte des Vorgehens beim Zusammenstellen der Getränkeprobe noch einmal enthält.

Appendix 2: Instructions in German as used in Study 4

First Instruction

Sie werden im Folgenden an einer Untersuchung teilnehmen, die den Zusammenhang zwischen Persönlichkeit und Geschmackspräferenzen untersucht. Ihre Antworten werden völlig anonym behandelt und Sie werden für die Dauer der Untersuchung mittels einer zweistelligen Nummer identifiziert, die auf all Ihrem Material stehen wird.

Im Rahmen von verschiedenen Studien werden unterschiedliche Typen von Nahrungsmitteln getestet. Daher möchte ich Sie zunächst bitten einen Fragebogen auszufüllen, in dem es um Ihre Geschmackspräferenzen geht.

Second instruction

Sie werden im Rahmen dieser Untersuchung einmalig eine Nahrungsmittelprobe kosten und diese bewerten. Die andere Versuchsperson im anderen Labor wählt diese für Sie aus.

Um Beeinflussung durch den Versuchsleiter zu vermeiden, werden Art und Menge der Proben von den Studienteilnehmern für einander ausgewählt. Sie werden daher im Folgenden eine Kostprobe für die andere Versuchsperson auswählen. Weil Menschen oft neugierig sind bezüglich der Geschmackspräferenzen der anderen Versuchsteilnehmer sehen Sie hier den Fragebogen zu den Geschmackspräferenzen der Person, für die Sie im Folgenden ein Getränk und dessen Menge auswählen sollen.

Zur gleichen Zeit wählt diese Person eine Probe für Sie aus. Die gewählten Proben werden dann ausgetauscht und jeder konsumiert und bewertet die jeweilige Probe. Danach ist die Studie beendet.

Die andere Versuchsperson wurde zufällig der Gruppe zugewiesen, die Getränke probiert. Aus einer Auswahl von drei Getränken wählen Sie ein Getränk und dessen Menge für die Versuchsperson im anderen Labor aus. Bitte probieren Sie die Getränke aus den dafür bereitgestellten Bechern, bevor Sie die Getränkeprobe für den anderen Versuchsteilnehmer wählen. Bitte mischen Sie die Getränke auf keinen Fall. Sondern wählen Sie bitte nur ein Getränk aus und geben eine Menge davon in den für die andere Versuchsperson vorgesehenen Becher.

Sie können die Menge des Getränks frei wählen. Sie können so viel oder so wenig des ausgewählten Getränks bereit stellen wie Sie möchten. Die Person soll die Getränkeprobe vollständig trinken.

Zusätzlich haben Sie auch ein Glas Wasser erhalten, so dass Sie mögliche Unannehmlichkeiten, die Sie eventuell durch das Probieren erleben, ausgleichen können. Dies ist nur für Sie zum Trinken. Mischen Sie auch das Wasser in keinem Fall mit dem Getränk für die andere Person.

Nachdem Sie ein Getränk und dessen Menge ausgewählt und es in den Becher gegeben haben, verschließen Sie den Becher mit der Alufolie. Dies dient dazu, dass

ich als Versuchsleiter nicht sehen kann, welches Getränk und wie viel davon Sie für die andere Person gewählt haben.

Bitte schreiben Sie dann die Nummer der anderen Versuchsperson auf den Becher. Ich werde dann den Becher mit dem gewählten Getränk der Versuchsperson im anderen Labor zum Trinken bringen. Hier erhalten Sie eine Checkliste, die alle Schritte des Vorgehens beim Zusammenstellen der Getränkeprobe noch einmal auflistet.

Ich werde in Kürze wieder kommen, um Ihre Getränkeprobe entgegenzunehmen und Ihnen einen Fragebogen und die Nahrungsmittelprobe, die momentan von der anderen Versuchsperson ausgewählt wird, überreichen

Erklärung

Erklärung gemäß § 8 Abs. 1 Buchst. b) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe.

Erklärung gemäß § 8 Abs. 1 Buchst. c) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Name, Vorname: Beier, Susanne

Heidelberg, 12. November 2012