# The Good, the Bad, and the Networks: Appetitive and Aversive Learning in the Balance of Mental Health

Laurens Tristan Kemp

ISBN: 978-94-6473-912-1

Illustrations and design: Louella Hendriksen

Print: Campusstore | Maastricht University

# The Good, the Bad, and the Networks: Appetitive and Aversive Learning in the Balance of Mental Health

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de Universiteit Maastricht,

op gezag van de Rector Magnificus, Prof. dr. Pamela Habibović

volgens het besluit van het College van Decanen,

te verdedigen in het openbaar op

donderdag 9 oktober 2025, om 13:00 uur

door

Laurens Tristan Kemp

**Promotores**

Prof. dr. Anita Jansen – Universiteit Maastricht

Prof. dr. Tom Smeets – Universiteit Tilburg

**Copromotor**

Dr. Katrijn Houben – Universiteit Maastricht

**Beoordelingscommissie**

Prof. dr. Harald Merckelbach (voorzitter) – Universiteit Maastricht

Dr. Pauline Dibbets – Universiteit Maastricht

Prof. dr. Peter de Jong – Rijksuniversiteit Groningen

Prof. dr. Reinout Wiers – Universiteit van Amsterdam

This PhD thesis is part of the project 'New Science of Mental Disorders' (www.nsmd.eu), supported by the Dutch Research Council and the Dutch Ministry of Education, Culture and Science (NWO gravitation grant number 024.004.016).

# Table of contents

# Chapter 1

General Introduction

**The Challenges in Understanding Mental Disorders**

A mental disorder is defined by the World Health Organization (WHO) as "a clinically significant disturbance in an individual's cognition, emotional regulation, or behaviour" (WHO, 2022). Common mental disorders encompass a diverse range of mood and anxiety disorders, such as depression and obsessive compulsive disorder (OCD; Stansfeld et al., 2016). However, mental disorders such as substance use disorders and externalizing disorders also have a substantial impact on mental health: a 2008 estimate indicates that between 18 and 36% of people worldwide suffer from a mental disorder in their lifetime (Kessler et al., 2009). The societal burden of mental disorders is comparable to cardiovascular and circulatory diseases, accounting for 13% of the global burden of disease in 2013 (Vigo et al., 2016). Moreover, mental disorders are responsible for premature mortality, such as through dementia and suicide, and this figure rose to 13 million excess deaths in 2010 (Charlson et al., 2015). In addition to this, comorbidity between mental disorders is typically high: someone diagnosed with a mental disorder is 14 times more likely to be diagnosed with another (McGrath et al., 2020).

Although mental disorders are an urgent problem (Patel et al., 2018), the current methods for treatment are rather underwhelming. Mental disorders are less frequently treated than other diseases: at least 35% of people worldwide who suffer from mental disorders do not receive treatment for them (WHO, 2011). Those that do receive treatment are more often given pharmaceutics than psychological therapies, despite that the latter is preferred by patients and is more effective in treating most mental disorders (Clark, 2018). Even so, such psychological treatments tend to lose their effectiveness over time, and relapse is common (Clark, 2018). The various challenges faced by psychological science to effectively treat mental disorders suggests that they are poorly understood (Cuijpers, 2019). Common symptoms of psychopathology are seen in multiple different types of mental disorders, which creates challenges for understanding the mechanisms by which mental health turns to disorder, without which prevention and treatment is substantially more difficult.

A recent approach to improving our understanding of mental disorders focuses on the interaction between psychological symptoms, which is known as the network perspective of psychopathology (Borsboom, 2017). This approach stands opposed to the categorical classification of mental disorders as if they were diseases that originate from the dysfunction of a particular part of the body or the brain. Instead, according to the network perspective, many mental disorders may not have an underlying cause at all, but rather the causal interplay between symptoms are the source of a mental disorder. Many people occasionally experience a bad night's sleep or a day of feeling under the weather but quickly recover from these symptoms without developing a mental disorder. However, if disturbed sleep and mood problems occur more frequently, these symptoms may fuel new symptoms such as fatigue, social isolation, and anhedonia, which may then reinforce each other even if the circumstances that caused them to occur have passed, resulting in depression. It is therefore hypothesized that a mental disorder is maintained by the interaction of all symptoms together, which prevent the patient from recovering from any individual symptom (Borsboom, 2017). From this perspective, it is easier to explain both the comorbidity of mental disorders and that they can be treated with the same interventions: treatment of the most common symptoms may help against many different disorders.

Although mutual reinforcement between symptoms is an important part of the network hypothesis, other aspects have also been argued to be crucial to understanding the role of network characteristics in mental disorder. One of these concerns the overall connectivity of an individual network, that is, how readily a change in one state such as mood can cause a change in another state such as sleep quality. This is argued to be a crucial factor in how vulnerable individuals are to developing a mental disorder: a person with strong network connectivity may easily shift into a state where multiple different symptoms are strongly activating one another, making them prone to develop a mental disorder as well as making it difficult to recover from it (Borsboom, 2017). A recent study testing this hypothesis did not find that individuals who score high on psychopathology have overall stronger connectivity than individuals  who score low (Jover Martínez et al., 2025). However, rather than showing overall stronger connectivity, it is possible that vulnerability to mental disorders is associated with specific parts of an individual's network of symptoms having higher connectivity. In other words, they may differ in their sensitivity to symptoms that are typical to a certain kind of disorder, such as internalizing versus externalizing disorders (Borsboom, 2017). There is currently little research available on whether individual differences are indeed associated with vulnerability to network effects that are seen in mental disorders. Testing this hypothesis may therefore improve our understanding of mental disorders and how well symptom networks represent them, but there are a great number of individual difference factors that could be associated with symptom networks, which necessitates careful consideration of which to study.

**Personality and Learning Models**

Personality traits are a type of individual difference factor that are commonly investigated for their association with mental disorders. Two important examples of these are the neuroticism and impulsivity personality traits. Neuroticism is strongly linked to common mental disorders (Kotov et al., 2010). Several models have been proposed to explain the complex and varied relationship between neuroticism and mental disorders. One such model is the spectrum model, which posits that neuroticism and common mental disorders are different manifestations of the same processes: all individuals have a certain level of neuroticism, but only in those with high levels does it rise to the severity of mental disorder (Ormel, Jeronimus, et al., 2013). Conversely, high impulsivity is associated with drug dependence, bipolar disorder and ADHD (Chamorro et al., 2012), suggesting that high levels of these traits also result in mental disorder, though the nature of this relationship is similarly complex, as is the case with neuroticism.

Personality traits are described as "consistent patterns in the way individuals behave, feel and think" (John & Gosling, 2000). One way that personality traits can be associated with mental disorders is when these consistent patterns are maladaptive and cause the individual stress. This could lead to mental disorders through mutually reinforcing effects between symptoms, as per the network hypothesis, but other mechanisms have also been proposed. Heinz et al. (2016) has suggested that investigating basic cognitive and behavioral processes, such as learning, and their relation to mental dysfunction is a more effective method than investigating mental disorders along their traditional diagnostic categories, given the wide variety of mental disorders that implicate dysfunctional learning. This has parallels with Reinforcement Sensitivity Theory (RST, Gray, 1975), which posits that different neural systems regulate adaptive behavior in response to competing goals: the

Behavioral Activation System (BAS), the Behavioral Inhibition System (BIS) and the Fight, Flight, Freeze System (FFFS) respectively. These are distinguished from one another by their function in driving the behavior appropriate for the situation: immediate danger (FFFS), (potential) reward signals (BAS) and mediation of conflict between the two (BIS; Pickering & Corr, 2008). These systems are theorized to cause the observed traits impulsivity (linked to BAS) and anxiety (linked to BIS). In other words, a person who scores high on impulsivity is thought to be more reactive to signals of reward, while a person who scores high on anxiety is thought to be more reactive to signals of punishment (Smillie et al., 2006). This theory also touches on the spectrum model: while these systems help someone to appropriately respond to their environment, disorders such as anxiety or panic disorder may be explained by the activation of the BIS at "maladaptive intensity" (Pickering & Corr, 2008). For example, a person with panic disorder may respond to even slightly ambiguous stimuli with strong BIS activation, that is, by evaluating them for signals of danger. An oversensitivity to bodily sensations that are associated with a panic attack could be a long-term outcome of this. If the spectrum model applies to both anxiety and impulsivity, then strong impulsivity may also be associated with an overactive BAS, which could result in maladaptive approach behavior, possibly leading to disorders such as addiction. It is also possible to model such interactions as time series effects in networks, which allows the network approach to test such hypotheses in addition to those specific to network structure, such as connectivity (Fried et al., 2017).

## Appetitive and Aversive Conditioning

Individuals' personality traits may thus predispose them to certain mental disorders, and especially sensitivity to reward and punishment are personality traits that may play a role in maladaptive behavior. One way of investigating this sensitivity to reward and punishment is via conditioning paradigms, which have been widely used both to compare patients with a mental disorder and healthy controls, and to compare measures of individual differences to performance on conditioning tasks. In conditioning paradigms, an initially neutral stimulus (the conditioned stimulus or CS) is paired with an intrinsically salient stimulus (the unconditioned stimulus or US) so that a participant learns to predict the occurrence of a US based on the presence of a

CS. When the US is a rewarding stimulus, this is known as appetitive reinforcement, and when it is a punishing stimulus, this is known as aversive reinforcement. CSs may consist of images or sounds and can range from simple (e.g. triangles, pure tones) to complex (e.g. faces, music), while USs are usually simple (e.g. electric shocks, loud noises) but may also be complex (e.g. erotic pictures, someone's favorite food). Learning is then measured through the conditioned response (CR) to the CS in anticipation of the US, which may consist of physiological measures such as skin conductance response (SCR) or behavioral measures such as food consumption. To determine whether learning has taken place, the CR to a CS that is followed by a US (the CS+) is compared to the CR for a CS that is not followed by a US (the CS-). For example, a conditioning task may display a triangle (CS+) followed by an electric shock (US) or a square (CS-) followed by no electric shock, and after several such displays, the experimenter can then compare the participant's response to the CS+ versus the CS-.

Although a variety of mental disorders are studied using conditioning paradigms, studies on fear and anxiety disorders are perhaps the most common. The hypothesis that anxiety disorder patients may have heightened responsiveness to aversive stimuli dates back almost half a century (Pitman & Orr, 1986). More recently, two meta-analyses have found that the most reliable result is that patients with anxiety disorder show worse discrimination between CS+ and CS- than healthy controls (Duits et al., 2015; Lissek et al., 2005). This result is primarily driven by increased fear responses to the CS-, showing that patients with anxiety disorder are more likely to respond fearfully to stimuli that do not signal danger, which may indicate impaired safety learning. Other mental disorders have also been investigated using aversive conditioning: in a study using functional magnetic resonance imaging (fMRI), patients who experience psychosis were shown to have lower activity in the ventromedial prefrontal cortex compared to healthy controls in response to neutral stimuli, which was interpreted as impaired safety learning (Quarmley et al., 2019). In men at risk for alcoholism, CS+/CS- discrimination was also impaired compared to healthy controls, but in contrast to anxiety disorder patients, their responses to the CS+ were weaker, suggesting reduced sensitivity to aversive learning (Finn et al., 1994). Aversive learning thus appears to play a role in different mental disorders, although it is difficult to draw conclusions on the relation of aversive learning to mental disorders other than anxiety given that these are less frequently studied.

In contrast to aversive conditioning, research into the relation between appetitive conditioning and mental disorders in humans is relatively uncommon. These studies focus primarily on eating disorders and obesity. Although obesity is not considered a mental disorder, increased food cue reactivity as a result of appetitive conditioning has been argued to play a role not only in obesity (van den Akker et al., 2018) but also in eating disorders like bulimia nervosa (Jansen et al., 1992). In both cases, strong appetitive CS-US associations are hypothesized to drive overeating. However, support for this hypothesis is inconsistent: studies comparing overweight and obese individuals to individuals with a healthy weight have shown poorer CS+/CS- discrimination (Coppin et al., 2014; van den Akker et al., 2017; Zhang et al., 2014), as well as better CS+/CS- discrimination (Meemken et al., 2018; Meyer et al., 2015) and no differences in discrimination ability (van den Akker et al., 2019). Other studies have investigated non-eating-related disorders: an fMRI study showed that schizophrenia patients performing an appetitive conditioning task showed stronger brain activation in response to the CS- compared to healthy controls, which suggested that they assigned more motivational salience to the CS- (Diaconescu et al., 2011). As with aversive learning, differences in appetitive learning may thus play a role in different mental disorders, highlighting the need for further study.

Moreover, studies that investigate both appetitive and aversive conditioning simultaneously and their relation to mental disorders are rare. According to RST, dysfunction in either appetitive or aversive learning can result in maladaptive behavior, but most studies only investigate one or the other type of learning process. Relevant exceptions include Zbozinek et al. (2021), who measured both aversive conditioning using electric shocks, and appetitive conditioning using monetary reward, in a sample of healthy volunteers. They found increased fear responses, that is, stronger aversive learning, in individuals with high trait anxiety, and lower expectations of monetary reward, that is, weaker appetitive learning, in individuals with high trait depression. In addition, Shook et al. (2007) used ambiguous stimuli that were associated with winning or losing points, which were used as appetitive and aversive stimuli. They found that depression and anxiety were associated with poorer appetitive learning, but not with differences in aversive learning. Other such studies have investigated depression (Huys et al., 2016) and PTSD (Elman et al., 2009), though not from an individual differences approach, which is comparatively understudied.

As the studies previously mentioned in this chapter have illustrated, appetitive and aversive learning are both relevant to a variety of mental disorders, suggesting that neglect of one or the other can leave critical insights undiscovered. However, and more importantly, sensitivity to appetitive and aversive learning may interact even in studies where only one of the two is measured. Gray (1975) posits that neutral stimuli may be perceived as either appetitive or aversive if they are contrasted with other stimuli. That is, in a conditioning task where an appetitive US is contrasted with the absence of a US, the latter may be perceived as aversive instead of neutral. In the opposite case, the absence of an aversive US may be perceived as appetitive. This suggests that differences in sensitivity to appetitive and aversive conditioning may in fact play a role in every conditioning study that compares a CS+ and CS-, and this possibility can be addressed by determining the association of both appetitive and aversive learning with symptoms of psychopathology.

**Learning Asymmetry: An Integrative Approach**

As mentioned in the previous section, Shook et al. (2007) investigated appetitive and aversive learning in a sample of healthy volunteers using a system of winning and losing points. Specifically, they compared the two by calculating the difference in how accurate each participant's responses were between appetitive trials and aversive trials, which they call *learning asymmetry*. They then compared this to self-report measures of depression and anxiety to determine if the relative strength of appetitive versus aversive learning had an effect. The approach of measuring learning asymmetry is notable for a few reasons. First, it is a measure of individual differences in sensitivity to appetitive or aversive learning. Conditioning research has been criticized for its shortcomings in the study of individual differences, partly because conditioning experiments tend to use simple, unambiguous stimuli to elicit strong conditioned responses, which are unsuitable for detecting between-person effects as differences between groups or conditions are more reliably detected when between-person variance is small (Beckers et al., 2013; Lissek et al., 2006). Second, it combines the measures of appetitive and aversive learning, creating a difference score indicating whether an individual is relatively more sensitive to appetitive or aversive learning, rather than simply showing stronger or weaker appetitive or aversive learning than other individuals. As mentioned previously, differences in conditioning effects between groups may be attributed to either appetitive or aversive learning, and the difference score can determine which of the two plays a greater role. Moreover, individual differences in general learning ability may result in some participants performing better at both appetitive and aversive trials than other participants, which the difference score corrects for.

While both Shook et al. (2007) and Zbozinek et al. (2021) studied appetitive and aversive learning and their relation to psychopathology, their methodology differed with regard to ensuring that they would be able to detect individual differences. Zbozinek et al. (2021) measured the effects of trait anxiety and trait depression on Pavlovian conditioning in an occasion setting paradigm, using both electric shocks and monetary reward. Simple colored figures (e.g. a blue triangle or a green star) were used as CS+ and CS-, while a trumpet sound or a violin sound were used as occasion setters to indicate whether it was possible for the US to occur at all. SCR and self-reports on fear and expectancy were measured. In this study, the reinforcement rate was 100%, that is, the CS+ was always followed by the US (given the right occasion setter). This study contains several elements typical to conditioning research: simple, unambiguous conditioned stimuli; learning measured by participants' responses to the CS+ versus the CS-; and separate experiments for appetitive and aversive conditioning. However, there are several weaknesses to this approach. Using colored figures as conditioned stimuli has the result that an attentive participant will typically memorize each possible outcome very quickly. This can be addressed by using a lower reinforcement rate: some studies present the US after the CS in only 80% of trials to add uncertainty to the outcome and avoid ceiling effects on expectancy ratings. Still, even in this case, learning only proceeds more slowly and the possible outcomes of each CS can be easily ascertained. This leads to what Lissek et al. (2006) termed *the strong situation*: a state of learning which leaves little ambiguity and makes it more difficult to detect variation between participants (see also Hedge et al., 2018). This ties into a broader discussion about contingency awareness and the experimenter demand effect, which question to what degree knowledge of the aims of the study influences participants' responses.

By contrast, Shook et al. (2007) used the Beanfest task, so named for the stimuli which are said to resemble beans. These stimuli consist of ovals containing dots, which vary in roundness and number of dots. Certain combinations of roundness and dot number are associated with either gaining or losing points, and their relation is not linear: some intermediate roundness and low dot number stimuli may be associated with gaining points, and some low roundness and high dot number stimuli may be associated with losing points. Participants are expected to learn which combinations of roundness and number of dots predicts point gain or loss, and must maximize their points by 'accepting' gain stimuli and 'rejecting' loss stimuli. This creates a set of highly ambiguous stimuli, which is advantageous for the detection of individual differences in learning. Participants received an additional monetary reward based on how many points they earned at the end of the task, and in this way the stimuli gained appetitive and aversive associations. On the one hand, this approach is further removed from the typical conditioning method of using primary reinforcers such as electric shocks, and it may result in weaker conditioned responses. On the other hand, contrasting learning from electric shocks with learning from monetary reward, as done by Zbozinek at al. (2021), presents the issue of comparing apples to oranges: the appetitiveness of monetary gain is qualitatively different from the aversiveness of pain (see also van der Schaaf et al., 2022). Comparing gaining points with losing points thus proportionally represents the relative strength of appetitive and aversive learning, though other pairings of appetitive and aversive stimuli are also possible. While this measure addresses several weaknesses of existing conditioning studies applied to the investigation of individual differences, it still leaves room for improvement. Stronger learning effects could be achieved with more distinct conditioned stimuli and more salient unconditioned stimuli, which should make it easier to detect individual differences in learning.

**Operationalizing Learning Asymmetry: Conditioned Stimuli and Confidence**

Given the above-mentioned considerations, we have developed a new task to measure learning asymmetry with the aim of making it as sensitive as possible to individual differences relevant to participants' learning and behavior. First, we used more distinctive conditioned stimuli, while maintaining their ambiguity. Although Beanfest used highly ambiguous stimuli that could be grouped together to improve the strength of their associations, their similarities strongly limited the degree to which they could be reliably learned. Instead, we used a variety of complex 3D objects as stimuli, which were designed by Watson et al. (2019) to be highly distinctive, varying across several stimulus dimensions including shape, color, and surface texture. On the one hand, each of these stimuli is sufficiently unique that they could be strongly associated with an appetitive or aversive outcome, as well as sufficiently abstract that existing associations would not interfere. On the other hand, they are sufficiently numerous that participants are unable to learn all associations perfectly. As a result, participants who are more sensitive to appetitive or aversive associations are expected to learn better from those stimuli associated with appetitive or aversive outcomes, effectively capturing their individual differences. As in Beanfest, these stimuli are presented up to four times, allowing participants to improve their learning over time, but unlike in Beanfest, two reinforcements are always equidistant, and the randomization process does not allow for two of the same stimuli to be presented close to each other. This prevents certain stimuli to be learned better due to their chance proximity to each other.

Second, confidence ratings were added to improve the observable variation in learning strength. Beanfest only compared response accuracy on identifying appetitive or aversive stimuli, but the reliability of this measure is weakened by the fact that participants have a 50% chance of responding correctly if they guess. This means learning asymmetry is more likely to show large variations by chance. To compensate for this, we measured confidence ratings in addition to accuracy. By asking participants how confident they are in their answer, we could obtain an additional measure of how strongly a particular association was learned. We can expect confidence to be low on correct guesses and downweight them in the calculation of the participants' learning asymmetry, and conversely, we can expect participants more sensitive to appetitive learning to form stronger associations with appetitive stimuli (and vice-versa) and thus give higher confidence ratings, increasing the likelihood that confidence-weighted learning asymmetry measures individual differences that truly exist.

## Aims of the Current Thesis

To summarize, a) network models provide an alternative method of studying mental disorders involving individual sensitivity to certain symptoms, b) individual differences in learning are related to personality traits and maladaptive behavior, c) previous research illustrates a variety of conditioning effects related to mental disorders, but lacks studies on both appetitive and aversive learning, and d) learning asymmetry is a method of investigating individual differences in learning that has rarely been applied to personality traits and mental disorders. In this thesis, the following empirical chapters present four studies that aimed to determine the relationship between learning asymmetry, personality traits and mental disorder symptoms. To this end, we developed a new learning asymmetry task and compared the associations of learning asymmetry with self-report measures on impulsivity, neuroticism, anhedonia, anxiety, substance use, and food reactivity. In addition, we determined whether the results on learning asymmetry and mental disorder symptoms are in line with the predictions following from RST and the network hypothesis. Finally, we investigated differences in day-to-day behavior of individuals who differ in learning asymmetry, to determine how sensitivity to appetitive versus aversive learning affects mood, impulsive behavior, avoidance and craving in a network model.

Chapter 2 describes the first version of the learning asymmetry measure using a novel task design incorporating a set of highly diverse but highly distinct stimuli, as well as a confidence measure intended to more accurately measure the strength of participants' learning. Both of these aspects are intended to create a conditioning task that maximizes variation between participants, which is necessary to detect effects caused by individual differences (Hedge et al., 2018; Lissek et al., 2006). In addition, a control version of the task was included that measures general learning rather than appetitive and aversive learning, to verify that the individual differences in learning asymmetry are indeed produced by variation in sensitivity to appetitive and aversive learning rather than variation of a different cause. Both tasks were compared to three self-report measures: the short form of the Eysenck Personality Scale – Revised, of which only the Neuroticism subscale was used; the UPPS-P Impulsive Behavior Scale, a measure of impulsivity comprised of five subscales, and the Domains of

Pleasure Scale, a measure of consummatory anhedonia. These self-report measures were selected to study traits that may each be associated with a different type of learning asymmetry: we hypothesized a positive relation between impulsivity and learning asymmetry, a negative relation between neuroticism and learning asymmetry, and a negative relation between anhedonia and learning asymmetry.

Chapter 3 builds on the previous study by using a broadly similar design for the learning asymmetry task yet used primary rather than secondary reinforcers. These primary reinforcers, namely sweet and bitter tastes, were intended to evoke stronger sensations than those experienced by participants in the previous study, and as such, make individual differences in learning easier to detect. Like the previous study, self-report measures of impulsivity and anhedonia were included, as were two self-report measures of mental disorder symptoms to determine their association with learning asymmetry: the Brief Symptom Inventory (BSI), a measure of psychological distress, and the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST), a measure of problematic substance use. These measures were used to investigate the association of the learning asymmetry with a wide variety of psychopathology symptoms, as well as to determine whether the results from the previous study could be replicated using a different type of reinforcer.

Chapter 4 investigates the theoretical significance of learning asymmetry by examining its relationship with RST. The traits central to RST are measured using the BIS/BAS scale, a self-report measure intended to capture individual variation in the Behavioral Inhibition System and Behavioral Activation System, which are posited to represent individual tendencies towards approach and avoidance behavior. Although the BAS is often described as representing impulsivity and the BIS as representing anxiety, these descriptions may not match with other traits with the same name. As learning asymmetry is a behavioral measure intended to capture individual differences, measuring its association with the BIS and BAS allowed us to determine whether behavioral and self-report measures of similar traits are compatible with one another. Furthermore, the predictions of RST can be used to determine whether appetitive and aversive learning as measured by our task are compatible with RST, or if they should be understood as different traits. Other self-report measures were included to clarify the learning asymmetry in both specific and general directions: the

Depression, Anxiety, and Stress Scale (DASS) to investigate mood disorder symptoms in further detail and the Power of Food Scale (PFS) to investigate food reactivity, an additional factor relevant to appetitive conditioning.

Chapter 5 explores the association between learning asymmetry and mood and behavior dynamics in participants' day-to-day lives, as measured by ecological momentary assessment (EMA), a data collection method which asks participants about their mental state and behavior multiple times throughout the day. This method is uniquely suited to discovering time-series effects, or how one's mental state and behavior is affected by the same at previous time points. These time series effects were examined through network analysis, a method which visualizes interactions between variables through nodes and edges. By comparing the networks of participants with either positive or negative learning asymmetry, we determined whether learning asymmetry is associated with differences in sensitivity to network effects that are potentially related to maladaptive behavior or symptoms of mental disorders, as predicted by the network hypothesis (Borsboom, 2017). The ways in which participants' reported mental state and behavior varies from one time point to the next, is illustrative of how individuals react to changes in mood and what drives changes in their behavior. The effect of learning asymmetry on these networks can therefore inform us of patterns in mood and behavior that are typical to individuals who are more or less sensitive to appetitive or aversive learning.

Finally, chapter 6 contains a general discussion of the results and conclusions from the four studies, examines the implications for learning asymmetry as a potential transdiagnostic factor in mental disorders, reflects on the limitations that are relevant to this research, and considers directions for future research.

# Chapter 2

Aversive Conditioning is Impaired in Impulsive Individuals: A Study on Learning Asymmetries

Published as: Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. (2024). Aversive conditioning is impaired in impulsive individuals: A study on learning asymmetries. *Journal of Behavior Therapy and Experimental Psychiatry*, *83*, 101939. https://doi.org/10.1016/j.jbtep.2023.101939

## Abstract

**Background and Objectives:** Appetitive and aversive conditioning are thought to be involved in the development and maintenance of mental disorders including anxiety, mood, eating, and substance use disorders. However, few studies measure the relative strength of appetitive and aversive associations, and their relevance to the risk of mental disorders. This study aims to address this gap.

**Methods:** We tested how readily healthy volunteers acquire appetitive vs. aversive associations. 150 participants associated complex 3D objects with either gain or loss and made decisions to gain or avoid losing points. We investigated the relationship of a learning asymmetry with neuroticism, impulsivity, and anhedonia, to test the hypothesis that a stronger learning asymmetry corresponds to more extreme scores on these traits.

**Results:** Impulsivity was positively associated with the learning asymmetry ($R^2$ = .10). This resulted from an inverse relation with the strength of aversive associations, indicating that impulsive individuals are worse at aversive learning. However, appetitive associations did not differ significantly. No correlations with neuroticism or anhedonia were found.

**Limitations:** Conditioning studies typically use primary reinforcers and a CS-. Lacking these may make these results less comparable to other studies.

**Conclusions:** We demonstrate that the learning asymmetry can measure individual differences linked to personality traits, and that impulsivity, normally linked with appetitive learning, also influences aversive learning. These results enable additional studies of learning asymmetry in relation to mental disorders, which could include measurements of mental health symptoms to provide further insight into how appetitive and aversive learning interacts with mental disorders.

## Introduction

Learning about the stimuli that predict appetitive and aversive outcomes is a skill that is indispensable for the thriving of animals, including humans. Feedback from our environment allows us to adjust our behavior to changing circumstances. However, an individual's learned behavior does not always result in optimal outcomes. Some may be predisposed to seek out appetitive outcomes more than they fear aversive ones, which may play a role in developing addictions (Albrecht et al., 2007). Others may be oversensitive to aversive outcomes and are unable to adapt their behavior towards seeking out appetitive outcomes even when this would be beneficial (Briere et al., 2010). Such maladaptive behaviors may be central to mental disorders such as depression, anxiety, and addiction. As such, investigating the causes and effects of these behaviors may help us understand these disorders.

The study of appetitive and aversive learning is frequently done using conditioning paradigms, which measure behavioral and/or physiological responses to repeated pairings of neutral stimuli with rewarding, appetitive stimuli or unpleasant, aversive stimuli. Aversive conditioning, that is, conditioning with stimuli causing fear or pain, has long been used to investigate whether anxiety disorders are associated with stronger acquisition and generalization of fear (e.g. Pitman & Orr, 1986). Results, however, are mixed, with studies showing greater fear acquisition among anxiety disorder patients for single-cue paradigms, but not more complex ones (Lissek et al., 2005). Generalization of fear was found to be broadened in PTSD patients (Kaczkurkin et al., 2017) but not in social anxiety disorder patients (e.g. Ahrens et al., 2016; also see Pittig et al., 2018 and Cooper et al., 2022, for reviews). It has yet to be explained why these results are inconsistent.

Appetitive conditioning studies, by contrast, are less common than those focused on aversive conditioning, partly due to the challenge of administering appetitive stimuli in a controlled setting. These have focused on behaviors associated with consumption, such as substance abuse and eating disorders (see Martin-Soelch et al., 2007 for a review). Other studies have used monetary reward in conditioning, showing that acquisition and extinction of the CS was associated with varying neural activity in the nucleus accumbens (Kruse et al., 2017). Furthermore, the acquisition and extinction of a monetary reward CS was found to be unrelated to intolerance of uncertainty (Morriss et al., 2021). However, impulsivity was shown to be related to increased food intake in an appetitive conditioning study using a virtual reality context CS (van den Akker et al., 2013a). Individual traits thus appear to be related to appetitive conditioning.

While aversive conditioning is often studied in relation to anxiety disorders and mood disorders, and appetitive conditioning in relation to substance and eating disorders, there are surprisingly few studies that have simultaneously examined both forms of learning. Most studies investigate absolute differences between patient groups and healthy controls in aversive or appetitive learning, but cannot inform us about potential relative differences in learning. Specifically, individual differences in the tendency to learn from aversive information relative to appetitive information may therefore reflect a transdiagnostic symptom that is shared by different mental disorders.

Most studies that have concurrently investigated appetitive and aversive learning have done so as part of fundamental conditioning research without addressing individual variability (Andreatta & Pauli, 2015; Baeyens et al., 1990; Gera et al., 2019; Kerkhof et al., 2011). Therefore, more insight may be gained by studying both forms of learning together, in combination with their relation to personality and other higher-level traits. The current study aims to do this by investigating the relative difference between how readily individuals acquire appetitive versus aversive associations, henceforth called the *learning asymmetry*. Our goal is to verify that the learning asymmetry is a meaningful measure of individual differences, by relating the learning asymmetry to various personality traits that have been linked to mental disorders.

**The learning asymmetry**

In this study, we used a novel behavioral task using Quaddles as conditioned stimuli: computer-generated 3D objects with multiple stimulus dimensions, designed to be highly distinctive, to facilitate learning and to reduce the likelihood that conditioned stimuli are mistaken for one another (M. R. Watson et al., 2019). This maximizes the associability of the conditioned stimuli and gives us the greatest chance of detecting the learning asymmetry. While conventional conditioning studies tend to use simple geometric figures (triangles, circles, etc.) as conditioned stimuli, these create highly unambiguous associations, which minimize the chance of finding variability between individual responses. For this reason, Beckers et al. (2013) and Lissek et al. (2006) argue that conditioning using 'weak' or ambiguous associations creates greater potential for the discovery of meaningful effects related to appetitive and aversive associations. We achieve this with a considerable number of Quaddles which should make each individual association relatively ambiguous, but given their uniqueness, allows them to remain distinguishable.

To explicitly compare the strength of appetitive and aversive associations, the choice of unconditioned stimuli is also important. Other studies have compared stimuli from different domains (e.g. painful electrical stimulation versus a small food reward; Andreatta & Pauli, 2015), but the issue here is that the appetitiveness of food is qualitatively different from the aversiveness of pain, making it difficult to compare these on the same scale (see also van der Schaaf et al., 2022). This can be addressed by using USs that are comparable in absolute value, such as using monetary gain and loss. Evidence suggests that how people learn and respond to these stimuli is in many ways similar to fear-inducing stimuli, including on measures of physiology (M. R. Delgado et al., 2006) and psychophysics (Schechtman et al., 2010). As these methods provide a comparable measure of conditioning, a system of point gain and loss was chosen as USs for this task, which was tied to an opportunity for a monetary reward in the form of additional study compensation. Thus, participants were incentivized to maximize their points, allowing these to serve as a secondary reinforcer.

**Personality traits**

To determine whether the learning asymmetry is meaningful as a measure of individual differences, we will relate it to existing measures of personality traits that are associated with mental disorders. Previous studies on personality traits linked to conditioning effects have included neuroticism, impulsivity, and anhedonia, which are typically involved in internalizing, externalizing and mood disorders respectively.

Neuroticism is defined as a trait sensitivity to aversive stimuli, and is strongly predictive of mental disorders including anxiety, mood and substance use disorders (Ormel, Bastiaansen, et al., 2013; D. Watson et al., 2005). Lommen et al. (2010) has shown that individuals high in neuroticism show more avoidance of ambiguous stimuli than those low in neuroticism in an aversive conditioning task. Additionally, a study of individual differences in fear acquisition showed impaired stimulus discrimination based on skin conductance responses for participants high in negative emotionality (Sjouwerman et al., 2020), a trait with substantial overlap with neuroticism (D. Watson et al., 2005). However, in a study comparing appetitive and aversive learning using 'weak' associations, Pietri et al., (2013) did not find an effect of neuroticism. This suggests that distinct factors in associative learning may be sensitive to variations in neuroticism, and due to its strong relation with common mental disorders, it is valuable to investigate this trait in relation to the learning asymmetry.

Impulsivity is a broadly defined trait that generally includes a propensity towards risk-taking behavior, poor impulse control, and difficulty maintaining attention. It is a diagnostic criterion for attention deficit and hyperactivity disorder (ADHD), antisocial personality disorder, and borderline personality disorder (American Psychological Association, 2013). Multiple studies show that high impulsivity is related to an increased attentional bias for appetitive stimuli in different contexts, including food reward (van den Akker et al., 2013a; Wardle et al., 2018a) and alcohol (Hicks et al., 2015a). In addition, impulsivity is also sometimes linked to impaired aversive learning (Patterson & Newman, 1993a; Wise & Dolan, 2020). In a review of studies on attentional bias and substance use, Field and Cox (2008) reviewed studies on attentional bias and substance use, and noted that the causal relationship between them is an important open question that remains to be determined. For these reasons, impulsivity is important to investigate in relationship to the learning asymmetry.

Anhedonia, namely a reduced or eliminated ability to find pleasure in different experiences, is a symptom common to several mental disorders, and core to schizophrenia and depression (American Psychiatric Association, 2013). Anhedonia has been implicated in the effects of reward expectation related to trait depression and anxiety (Zbozinek et al., 2021) and as a risk factor for schizophrenia (Kwapil, 1998; Moran et al., 2022). In a study comparing different types of reinforcement, Banica et al. (2022) found significant associations between neural activity waveforms and self-reports of consummatory pleasure and social anhedonia, but only for tasks using food and social reward, respectively. By contrast, the monetary reward task did not show an association with self-reported anhedonia. Furthermore, multiple studies using the Probabilistic Reward Task have indicated that anhedonic symptoms were inversely related to a response bias towards a more frequently rewarded stimulus (see Kangas et al., 2022 for an overview). These suggest anhedonia is likely to show an association with the learning asymmetry.

As a first step to examining whether the learning asymmetry is indeed relevant to mental disorders, the present study sought to investigate whether learning asymmetry is meaningfully related to the personality traits neuroticism, impulsivity and anhedonia that have been linked to mental disorders. We hypothesized that neuroticism, impulsivity and/or anhedonia would be linearly or inversely associated with the learning asymmetry. This would demonstrate that the learning asymmetry is a meaningful measure of individual differences and has the potential to provide more insight into the relations between conditioning, personality traits, and mental health. Additionally, we expected successful appetitive and aversive conditioning to be reflected in participants' responses, such that repeated reinforcements would increase their accuracy and confidence at identifying stimuli that led to appetitive and aversive outcomes, similar to that reported by Shook et al., (2007).

# Method

## Participants

150 participants provided informed consent and completed the experiment. Recruitment was done via physical and online advertisements in a variety of spaces, but respondents primarily consisted of undergraduate students (74.66%). Participants' nationality was predominantly German (40%) and Dutch (30.66%), with other European nationalities making up 18% of the sample, non-European nationalities 7.33%, and 4% reporting multiple nationalities. Participants were between 16 and 35 years of age (*M* = 21.6, *SD* = 3.36), and consisted of 31 males, 119 females. Participation occurred between 9 November 2021 and 17 February 2022. Participants took part in the study for either study credit or a €5 voucher as compensation.

Furthermore, participants who scored above average on the learning task were given a €5 voucher in addition to their chosen compensation.

**Materials**

Stimuli were generated using the scripts and instructions provided by Watson et al. (2019) on 3DS Max 2022. Two sets of twenty stimuli were generated, and ten stimuli of each set were assigned to each trial type. For the experimental task, they were given valenced associations: ten resulted in point gain, and ten in point loss. For the control task, they were given unvalenced associations: ten were associated with the Seeds category, and ten with the Spores category. These associations were counterbalanced between participants, as were the sets used for the respective task, and the order in which the experimental and control tasks were performed.

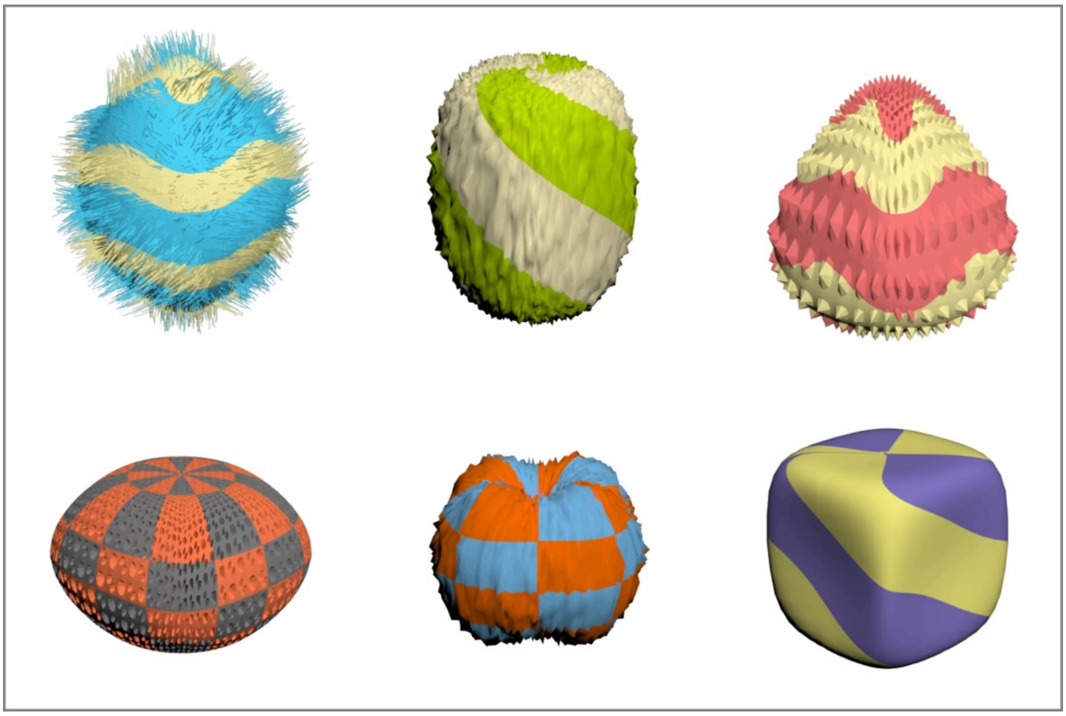


*Figure 1: Some examples of Quaddles used as conditioned stimuli.*

### *Task Performance*

After completing the task, participants were asked nine questions about various aspects of the task, to verify that they completed the task as intended and that no problems occurred with the online test platform. Details on these questions are available in Supplementary Table 1.

### *Eysenck Personality Questionnaire*

To assess neuroticism, the Neuroticism subscale of the short form of the Eysenck Personality Questionnaire, Revised (EPQ-R) was obtained (Eysenck & Eysenck, 1975). This questionnaire measures the three main dimensions of personality. Aside from Neuroticism, the EPQ-R also measures the Extraversion and Psychoticism subscales, and a Lie scale which aims to determine the respondent's truthfulness. It consists of 48 Yes/No questions, 12 of which belonged to the Neuroticism subscale. To obtain the score for this subscale, all 'Yes' answers were summed. Cronbach's alpha for this subscale was high ($\alpha = .83$). The remaining subscales were not included in this report.

### *UPPS-P Impulsive Behavior Scale*

The UPPS-P Impulsive Behavior Scale is an inventory of various measures of impulsive behavior, modeled after five factors that were assessed by exploratory and confirmatory factor analysis (Lynam et al., 2006). It is named after its subscales, which include (Negative) Urgency, (lack of) Premeditation, (lack of) Perseverance, Sensation Seeking, and Positive Urgency. It consists of 59 items scored on a scale of 1 (Agree Strongly) to 4 (Disagree Strongly). To obtain the total impulsivity score, the average score of each subscale was calculated, then all subscales were averaged. Cronbach's alpha for the entire scale was very high ($\alpha = .93$).

### *Domains of Pleasure Scale*

The Domains of Pleasure Scale (DOPS) is a measure of consummatory anhedonia, or the (lack of) pleasure obtained from performing various activities

(Masselink et al., 2019). Its subscales include the social, sexual, perceptual, and personal achievement domains. It consists of 21 items scored on a 100-point visual analog scale, on which the left extreme represents 'Not at all' and the right extreme represents 'Very much'. To obtain the total anhedonia score, the average score of each subscale was calculated, then all subscales were averaged. Cronbach's alpha for the entire scale was high ($\alpha = .88$).

**Procedure**

The experimental procedure was approved by the local Ethics Review Committee Psychology and Neuroscience. Participants were invited to perform the experiment using their own computer or mobile device. They were provided information about the study and gave informed consent via an online form provided via Qualtrics. After this, the experiment was run using the Inquisit online platform, version 6.5.0. Participants were provided with on-screen written instructions, followed by six practice trials, followed by either the experimental task or the control task.

In each task, the first block was dedicated to teaching the participant the association of each stimulus, and no responses were necessary. This will be referred to as the acquisition block. The subsequent three blocks asked the participant to respond to each stimulus based on the associations they learned from the acquisition block. These will be referred to as the response blocks. After each block, participants were allowed to pause for as long as they wished, during which they also had the opportunity to re-read the instructions.

First, the procedure for the experimental task will be described. The first block, or acquisition block, indicated the association of each stimulus ("Gains 10 points"/"Loses 10 points"). During the subsequent three blocks, or response blocks, the participant was asked to 'accept' or 'reject' the outcome of each stimulus. The goal was to gain points and avoid losing points by 'accepting' stimuli associated with point gain and 'rejecting' stimuli associated with point loss. This was done by clicking one of two buttons, 'Yes' or 'No', when given the question "Accept the result?". On trials where the stimulus signaled point gain (hereafter: 'Gain trials'), responding 'Yes' would add 10 points. Conversely, on trials where the stimulus signaled point loss (hereafter: ‘Loss trials’), responding ‘Yes’ would subtract 10 points. Responding ‘No’ resulted in neither point gain nor loss. In this way, Gain trials created appetitive associations through the possibility of gaining points, and Loss trials created aversive associations through the possibility of losing points.

After responding, participants selected a point on a 100 point visual analog scale to indicate their confidence in their decision. They then received feedback on how their points changed, if at all. In addition, the participant’s point total for the entire task was displayed with each instance of feedback. For more details, see Figure 2. The procedure for the control task was mostly identical, but used non-valenced associations instead, namely ‘Seeds’ and ‘Spores’ in place of Gain and Loss. Correct answers would add 10 points to the participant’s point total, while incorrect answers would result in no change. In this way, the control task served as a baseline comparison to the experimental task, allowing us to determine whether any detected effects would be caused by the valence asymmetry or if they resulted purely from the method of learning about two mutually exclusive categories.

Upon completion of both the experimental task and the control task, participants answered the task-related questions, and filled out the EPQ-R, the UPPS-P, and the DOPS. Participants were thanked and fully debriefed at the end of the study and received a gift voucher or course credit as remuneration for their participation in the study.

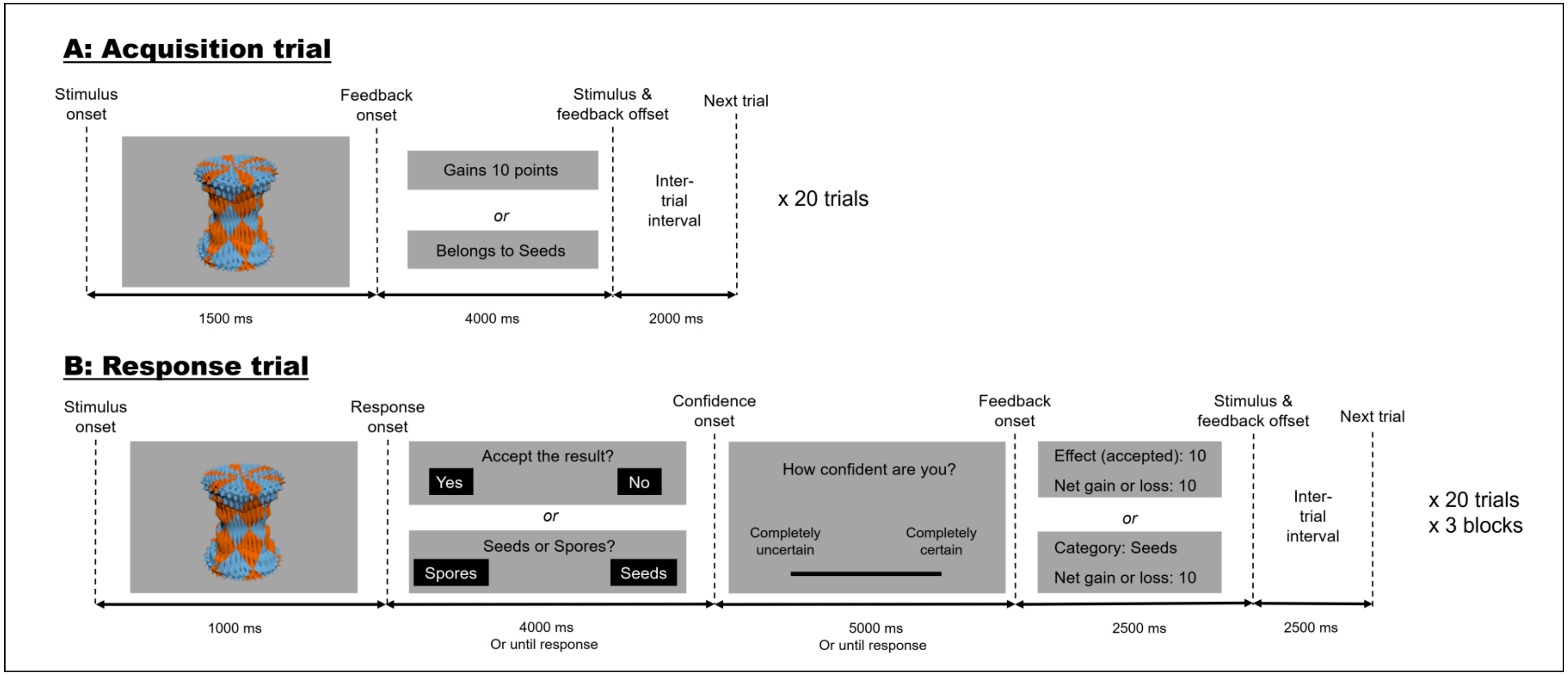



*Figure 2: Diagram of the task procedure for Acquisition trials (A) and Response trials (B). Areas in grey represent elements displayed during the task. For both trial types, if two elements are shown, those on the top represent options and feedback given on the experimental task, and those on the bottom represent options and feedback given on the control task. For the former, the figure shows example feedback for a Gain stimulus and a Yes response, and for the latter, a Seeds stimulus and a Seeds response.*

## Outcome measures

From the measures of accuracy and confidence, a composite variable was created using the sum of confidence ratings on all correct trials minus the sum of confidence ratings on all incorrect trials, then divided by the total number of trials, a value that will henceforth be referred to as 'calibration'. By weighting participants' answers according to their confidence, calibration more closely represents the strength of the learned associations.

## Transparency and Openness

The study's hypotheses, analyses, exclusion criteria and sample size were preregistered at https://aspredicted.org/vv6cq.pdf. All data, analysis code, stimulus materials, and task script are available at https://osf.io/cgqna/. This report was written following the standards of JARS (Kazak, 2018). Using G*Power version 3.1.9.7 (Faul et al., 2007) for an a priori power analysis, it was determined that to achieve a power of 0.95 for a medium effect at a significance criterion of $\alpha = .05$, a minimum sample size of 119 would be required for a linear multiple regression analysis with 3 predictors. Given that the reliability of behavioral data obtained online may be lower, sample size was set to 150 participants to compensate. Furthermore, given the high dropout rate for an online study compared to a lab study, only the first 150 participants who fully completed the experiment were included in the sample.

Exclusion criteria were specified as follows: Failure to respond within the time limit on 15 trials or more; accuracy not significantly different from chance, and their calibration score below 5 across all blocks; questionnaire answers indicating poor effort or the involvement of automation; or a score differing more than three standard deviations from the mean on the EPQ Lie scale. Additionally, participants were considered outliers if their data showed a difference greater than 2.75 standard deviations from the mean on the following variables: response bias, accuracy asymmetry, and calibration asymmetry.[1] Response bias represents the difference between the proportion of participants' responses to each type of trial and the actual

[1] In a normally distributed sample of 150, the expected number of values beyond 2.75 standard deviations is less than one, implying that values that fall outside this range did not originate from the same distribution.

proportion of those trials. Outliers on the response bias were excluded, and outliers on the accuracy asymmetry and calibration asymmetry were recoded to the closest non-outlier value.

## Results

All the following analyses were performed in IBM SPSS Statistics version 26.

### Inclusion and Exclusion

Based on the criteria for response bias, six participants were excluded. Based on the averages across all blocks, six participants were considered outliers on the learning asymmetry and their values were recoded to the closest non-outlier value, and based on the averages per block, five participants had their values recoded. These changes did not affect the significance of the results. No participants met the other listed exclusion criteria, resulting in a final sample size of 144 participants (115 female, 29 male).

### Task Performance

Initial tests aimed to verify that participants showed a learning effect and that their learning was reflected in the outcome variables. First, confidence ratings as an indicator of learning were tested. Paired-samples t-tests (two-tailed) were performed on the average confidence rating values for all correct trials versus all incorrect trials, for both the experimental task (exp) and the control task (ctrl). Results showed that participants had significantly higher confidence ratings on trials they answered correctly, $M_{exp} = 61.91$, $SD_{exp} = 17.41$, $M_{ctrl} = 57.11$, $SD_{ctrl} = 19.01$, compared to trials they answered incorrectly, $M_{\text{exp}} = 47.89$, $SD_{\text{exp}} = 17.85$, $M_{\text{ctrl}} = 46.75$, $SD_{ctrl} = 18.46$. This was true for both the experimental task, $t(143) = 13.14$, $p < .001$, and the control task, $t(143)= 10.84$, $p < .001$. This confirms that participants were able to accurately introspect about their own performance, providing additional information about the strength of their learning.

Second, participants' performance over time was tested by performing two repeated-measures analyses of variance (ANOVAs) using Condition (experimental vs.

control) and Block (1 vs. 2 vs. 3) as factors and accuracy and calibration as the dependent variables. The accuracy tests showed significant main effects of Condition, $F(1,143) = 13.96$, $p < .001$, and Block, $F(2,142) = 36.90$, $p < .001$, but no significant interaction, $F(2,142) = 1.01$, $p = .73$. Paired t-tests showed that the increase in accuracy between block 1, $M = .64$, $SD = .12$, and block 2, $M = .65$, $SD = .12$, was not significant, $t(143) = -1.62$, $p = .22$, but the increase between block 2 and block 3, $M = .70$, $SD = .13$, was significant, $t(143) = -7.20$, $p < .001$.

The calibration tests showed similar results, namely significant effects of Condition, $F(1,143) = 17.97$, $p < .001$, and Block, $F(2,142) = 60.50$, $p < .001$, and no interaction effect, $F(2,142) = 1.38$, $p = .51$. Paired-samples t-tests showed that the increase in calibration between block 1, $M = 20.43$, $SD = 1.35$, and block 2, $M = 18.76$, $SD = 1.56$, was significant, $t(143) = -2.78$, $p = .012$. The increase from block 2 to block 3, $M = 32.42$, $SD = 1.83$, was also significant, $t(143) = -9.67$, $p < .001$. These data confirm that the calibration measure accurately represents participants' learning over time, and that their improved learning was reflected in increased confidence ratings, even when this was not detectable in their accuracy. Specific measures of task performance per block can be found in Table 1.

Finally, the development of task performance over time was investigated by comparing Gain and Loss trials across three consecutive response blocks. Differences between blocks were calculated using two repeated-measures ANOVAs using Block and Valence as factors, and with one analysis using accuracy as the dependent variable, and the other using calibration. The assumption of sphericity was violated for both accuracy, $\chi^2(2) = 6.45$, $p = .040$, and calibration, $\chi^2(2) = 14.96$, $p = .001$, and so the Greenhouse-Geisser correction was applied. For accuracy, results showed a significant effect of Block, $F(2,142) = 22.12$, $p < .001$, as well as an interaction between Block and Valence, $F(2,142) = 3.75$, $p = .026$, but no main effect of Valence, $F(1,143) = 2.23$, $p = .138$. For calibration, results showed a significant effect of Block, $F(2,142) = 47.99$, $p < .001$, but no significant effect of Valence, $F(1,143) = 1.60$, $p = .21$, and no significant interaction effect, $F(2,142) = 1.84$, $p = .16$. The interaction effect on accuracy was driven by significantly higher accuracy on Loss trials in block 3, $M = .76$, $SD = .20$, than Gain trials, $M = .70$, $SD = .20$, $t(143) = 3.241$, $p = .001$. Differences in accuracy on blocks 1 and 2 were not significant (all *p*s > .05).



.

Additional analyses investigating task performance effects can be found in the supplementary material.

**Table 1**

*Task Performance Descriptive Statistics for the Experimental and Control Tasks*

| Trial type | Variable | Block 1 | | Block 2 | | Block 3 | |
|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Experimental | | | | | | | |
| Gain trials | Accuracy | .65 | .19 | .67 | .18 | .70 | .20 |
| | Calibration | 22.61 | 26.49 | 26.46 | 27.35 | 33.64 | 32.32 |
| Loss trials | Accuracy | .64 | .23 | .66 | .23 | .75 | .21 |
| | Calibration | 22.10 | 29.53 | 26.08 | 30.80 | 39.09 | 30.14 |
| Control | | | | | | | |
| Seeds trials | Accuracy | .60 | .18 | .61 | .18 | .68 | .19 |
| | Calibration | 13.59 | 23.86 | 15.72 | 26.29 | 25.18 | 28.56 |
| Spores trials | Accuracy | .64 | .17 | .63 | .18 | .67 | .20 |
| | Calibration | 20.59 | 23.95 | 21.47 | 25.39 | 29.83 | 28.24 |

Note: Higher calibration values represent a greater ability to respond correctly and with confidence in correct answers.

### Learning Asymmetry and Personality Measures

The main hypothesis was tested by a linear multiple regression analysis using the total scores on the UPPS-P, DOPS, and EPQ-N as predictors separately for the accuracy asymmetry and the calibration asymmetry. A backwards entry method was used for these predictors, with the criterion of probability *F*-to-Remove >= 0.1. As shown in Table 2, only the models for the experimental task, but not the models for the control task, explained a significant amount of the variance in the accuracy asymmetry and the calibration asymmetry. Furthermore, both Neuroticism and Anhedonia did not exceed the removal criterion, indicating they had no significant prediction effect in the model. In other words, only Impulsivity was left as a significant predictor of the learning asymmetry. On the control task, the three predictors did not account for a significant part of the variance in the accuracy asymmetry or the calibration asymmetry. The following subsections will unpack the effect of Impulsivity in post-hoc analyses on the relative contribution of Gain and Loss trials, and the UPPS-P subscales.

**Table 2**

*Linear Regression Analyses for Experimental and Control Asymmetry Variables*

| | Experimental task | | Control task | |
|---|---|---|---|---|
| | Accuracy Asymmetry | Calibration Asymmetry | Accuracy Asymmetry | Calibration Asymmetry |
| Included | All | All | All | All |
| $F_{change}(3,140)$ | 7.748 | 5.297 | 2.264 | 1.800 |
| *p* | **<.001** | **.002** | .084 | .150 |
| $R^2$ | .142 | .102 | .046 | .037 |
| Adjusted $R^2$ | .124 | .083 | .026 | .017 |
| Included | Impulsivity, Anhedonia | Impulsivity, Anhedonia | Neuroticism, Anhedonia | Impulsivity, Anhedonia |
| $F_{change}(1,140)$ | 0.122 | 0.058 | 2.417 | 0.596 |
| *p* | .728 | .811 | .122 | .441 |
| $R^2$ | .142 | .102 | .030 | .033 |
| Adjusted $R^2$ | .129 | .089 | .016 | .019 |
| Included | Impulsivity | Impulsivity | Neuroticism | Impulsivity |
| $F_{change}(1,141)$ | 0.001 | 0.400 | 0.035 | 0.477 |
| *p* | .974 | .528 | .851 | .491 |
| $R^2$ | .142 | .099 | .030 | .030 |
| Adjusted $R^2$ | .136 | .093 | .023 | .023 |

Note: Each column represents one regression analysis. The cells under 'Included: All' denote the test statistics for the model with all predictors (Impulsivity, Anhedonia and Neuroticism) included. The remaining cells under 'Included' denote the change in test statistics for the model when the predictor that is not listed is excluded due to failure to exceed the removal criterion. Non-significant changes resulting from this exclusion indicate that the predictor did not significantly affect the model. Bolded cells indicate *p*-values under 0.05.



.

### *Gain versus Loss trials*

The first step in unpacking the positive association with Impulsivity was comparing different trial types. Separate Pearson correlation coefficients (two-tailed) were calculated for the accuracy and calibration on Gain and Loss trials, and Impulsivity. Gain trials did not show significant correlations with Impulsivity on either accuracy, $r(142) = .10$, $p = .213$, or calibration, $r(142) = .04$, $p = .65$. However, Loss trials showed significant correlations with Impulsivity on accuracy, $r(142) = .32$, $p < .001$, and calibration, $r(142) = -.27$, $p = .001$, indicating that high impulsive individuals performed worse on these trials, while low impulsive individuals performed better.

### *UPPS-P Subscales*

The second step in unpacking the effect of Impulsivity on the learning asymmetry was to compare the different subscales of the UPPS-P. The subscales in question are known as Positive Urgency, Negative Urgency, (lack of) Premeditation, (lack of) Perseverance, and Sensation Seeking. Two linear multiple regression analyses using these subscales as predictors and the calibration asymmetry on both tasks as the dependent variable were performed using the backwards entry method. As expected based on the previous result, the model including all five predictors explained a significant amount of the variance in the value of the calibration asymmetry, $F_{change}(5,132) = 4.24$, $p = .001$, $R^2 = .13$, $R^2_{Adjusted} = .10$. All predictors excluding Sensation Seeking and Positive Urgency were excluded from the model as they did not exceed the removal criterion, which did not lead to any significant changes to the model, all $F_{change} < 1$, all $p$s > .326, all $R^2_{change} < .01$, leaving Sensation Seeking and Positive Urgency as the only significant predictors, $F(1,136) = 10.19$, $p < .001$, $R^2 = .13$, $R^2_{Adjusted} = .11$.

## Discussion

In this study, we developed a conditioning task with the aim of assessing individual differences in how easily appetitive versus aversive associations are learned. Participants performed an appetitive and aversive learning task, and we calculated an individual learning asymmetry and investigated its relationship with three personality traits that are associated with different mental disorders. Main effects of learning were significant, and participants were able to identify experimental and control stimuli better with each learning block, with accuracy on Loss stimuli outpacing Gain stimuli after three reinforcements. This can be attributed to the well-known phenomenon of loss aversion (e.g. Sokol-Hessner & Rutledge, 2019), which may have resulted in increased motivation to remember Loss trials. In short, participants showed reliable appetitive and aversive learning.

Results showed that impulsivity was a significant predictor of the learning asymmetry, in that higher impulsivity was associated with worse learning of aversive associations, while appetitive learning was unrelated to impulsivity. As impulsivity is generally shown to be associated with attentional bias in appetitive conditioning studies (e.g. Hicks et al., 2015; van den Akker et al., 2013; Wardle et al., 2018), our finding that aversive, but not appetitive learning is linked to impulsivity is surprising. However, to our knowledge, no other studies have investigated impulsivity in the context of both appetitive and aversive learning – although some have posited a link between impulsivity and impaired aversive learning (Patterson & Newman, 1993a; Wise & Dolan, 2020) – meaning comparisons to other studies may be difficult. The current results suggest that the focus on impulsivity in relation to appetitive learning is insufficient to gain a full understanding of its relevance to mental disorders related to appetitive and aversive learning. Instead, taking a broader view of individual differences in associative learning and certain maladaptive patterns linked to impulsivity, it may be possible to use prevention and intervention tailored to the individual, as supported by Lonsdorf and Merz (2017).

Our study also included two null results, namely for neuroticism and anhedonia. We expected a possible association of learning asymmetry with neuroticism, as it is strongly associated with common mental disorders (Ormel, Bastiaansen, et al., 2013). However, no such relation was found, which suggests it



.

may not play a role in participants' propensity to acquire appetitive versus aversive associations. Although it was negatively correlated with task performance in general (see pp. 39), this influence does not correspond to a relative difference between appetitive and aversive learning, meaning it may simply affect performance through impaired attention or motivation. Other traits related to aversive learning, such as intolerance of uncertainty (Morriss et al., 2021), may be more relevant to investigate in future research on the learning asymmetry.

The lack of an association of the learning asymmetry with anhedonia may be attributable to the fact that the current study used a monetary incentive, rather than a different type of reward. Banica et al. (2022) showed a correlation between consummatory pleasure, which the DOPS measures, and the neural activity waveform associated with a food reward, but not a monetary reward. This suggests that using a different reinforcer has the potential to reveal any behavioral effects of anhedonia that a monetary reward could not.

A major limitation of the current study is the lack of stimuli that are not paired with any outcome (i.e. a CS-), which was a compromise for several reasons: the duration of the task, the number of stimulus exposures per participant, and the cognitive load required for the participants to consider three possibilities rather than two. Although we have compensated for this by using a non-valenced control task, it remains difficult to compare the results between tasks. Our conclusions may therefore not fully generalize to other conditioning studies, and future research should attempt to replicate current findings using a task that includes both CS+ as well as CS-.

The original aim of this study was to develop a learning task that could accurately measure participants' propensity to acquire appetitive and aversive associations, and to link this measure to personality traits relevant to mental disorders. The evidence suggests the current task fits this objective, as shown by the clear differences between the experimental and control tasks, as well as overall task performance and the finding of a differential effect of impulsivity. However, the applicability of these results to the study of mental disorders is limited due to a lack of data on their symptoms, which was not gathered in favor of using questionnaires that would provide a personality scale regardless of a participant's mental health. With the current study's evidence that the learning asymmetry is a valid measure of individual difference, future research could gain additional insight to its role in mental disorders by investigating its relationship with different symptoms.

In conclusion, the learning asymmetry, a novel measure of relative learning ability, was shown here to have a significant association with impulsivity, opening the door for further research using this measure in relation to personality and mental disorders. These results indicate that measuring individual differences in both appetitive and aversive associations, rather than either in isolation, has the potential to provide greater insight into how impulsivity, and likely other personality traits, affect learning behavior. Given the preponderance of studies relating common mental disorders with either appetitive or aversive learning, further research incorporating both of these, as well as measuring mental health symptoms, is recommended.

## Supplementary Material

**Supplementary Table 1**

*Questions and Answers from the Task Follow-Up Questionnaire*

| Question | Answers |
|---|---|
| 1. What did you use to perform the task? | • A mouse<br>• A touchpad or something else |
| 2. Were the tasks and instructions clear and understandable? | a) Yes, I understood everything perfectly<br>b) Yes, I understood most of it<br>c) No, there were a lot of things I didn't understand<br>d) No, I didn't understand any of it |
| 3. Please elaborate on what you found unclear or confusing, if you can. | |
| 4. Did you experience any technical problems or real-life interruptions during the tasks? If yes, please describe what happened. If no, you can leave this blank. | |
| 5. Did you put a good effort into giving accurate answers? | a) Yes, I did the best that I could<br>b) Kind of, I didn't try as hard as I could have<br>c) Not really, I wasn't paying much attention<br>d) No, I was just clicking randomly |
| 6. Which of these words match your feelings about the task? Please select all that apply. | ▪ Interesting ▪ Boring<br>▪ Frustrating ▪ Difficult<br>▪ Challenging ▪ Rewarding |

| | |
|---|---|
| 7. How distinguishable did you find the different objects? | a) They were all clearly different<br>b) They were somewhat similar but still distinguishable<br>c) They had a lot of similarities and I mixed some of them up<br>d) They were so similar I mixed them up constantly |
| 8. Did you use any memorization techniques or other methods that helped you respond accurately? If yes, please describe them. If no, you can leave this blank. | |
| 9. What did you think the purpose of the task was? If you have not thought about this, please give your best guess. | |

Note: Empty cells in the 'Answers' column indicate that participants were allowed to type in their own answer.

In the task-related follow-up questions, 97.3% of participants indicated that the instructions were mostly or fully understandable and 99.3% indicated making at least a moderate effort (Table 2). When asked about the distinguishability of the stimuli, 54.6% of participants reported them to be mostly or very distinguishable, which is expected given that the stimuli consisted of many combinations of a few different properties. As such, these answers indicated that the balance between recognizability and difficulty was well-struck.

**Supplementary Table 2**

*Percentage of answers to each of the multiple choice questions regarding the learning task. Response levels represent an approximation of the multiple-choice answer. For the exact phrasing, see Table 1.*

| Question | Answer | | | |
|---|---|---|---|---|
| | Yes, very much. | Yes, mostly. | No, not very. | No, not at all. |
| Were the instructions clear and understandable? | 75.3% | 22% | 2.7% | 0% |
| Did you put a good effort into giving accurate answers? | 79.3% | 20% | 0.7% | 0% |
| Did you find the different objects distinguishable? | 8% | 46.6% | 40% | 5.3% |

The answers provided to the open questions indicated no recurrent issues with the task, with one exception: some participants reported using a strategy in which they deliberately focused on only one type of association and tried to ignore the other. For example, some reported that they paid attention to the stimulus whenever a Gain trial occurred, and looked away from the screen whenever a Loss trial occurred. This allowed them to identify the associations based on familiarity rather than by the strength of the associations themselves. Using this strategy distorts the learning asymmetry for these participants, as their performance is no longer tied to their learning ability, but rather what they choose to focus on.

Seven participants reported using this strategy, so their calibration asymmetry values were compared to the rest of the sample by a one-way ANOVA. This indicated that their calibration asymmetry was not significantly different from the mean, $F(1,136) = 1.56$, $p = .214$. Participants who reported using this strategy had a slightly more positive calibration asymmetry, $M = 8.6$, $SD = 15.02$, than participants who did not, $M = -1.89$, $SD = 21.93$. Based on this result, these participants were not excluded from the analysis.

Other reported strategies included verbal repetition, memorizing specific stimulus features, attempting to determine rules by which the result could be predicted, assigning the stimuli names or other identifiers, and memorizing the sequence. However, the experimental design did not allow for any of these strategies to distort the outcome measures, and so no action was taken on the basis of participants' reported strategies.

**Task performance**

In addition to the analyses reported in pp. 28-29, we tested whether the learning asymmetry was confounded by participants' overall accuracy on the task. It is possible that participants with particularly high or low accuracy could also have a larger or smaller learning asymmetry. To investigate this, accuracy asymmetry and calibration asymmetry were converted to their absolute values, and were respectively correlated with the total accuracy and total calibration on their respective tasks. The absolute values were used so that the size of the asymmetry in both the positive and negative directions could be compared to the total accuracy and calibration. This measured to what degree the learning asymmetry was caused by differences between appetitive and aversive learning, rather than differences in general learning ability. As this transformation resulted in a non-normal distribution, Spearman's rho (one-tailed) was used to estimate non-parametric correlations. This revealed significant negative correlations for the accuracy and its asymmetry on both the experimental task, $r(136) = -.26$, $p = .001$, and the control task, $r(136) = -.25$, $p = .036$. However, the correlations were negligible for the calibration and its asymmetry on the experimental task, $r(136) = -.06$, $p = .250$ and the control task, $r(136) = .02$, $p = .403$. Furthermore, both correlations with accuracy were negative, meaning that participants who best identified Gain and Loss stimuli also showed the lowest degree of asymmetry, whether in the positive or the negative direction. This is likely due to a ceiling effect, as the more accurate a participant's responses, the less variation can occur between their Gain and Loss accuracy. As a result, it is more difficult to determine whether these participants more readily acquired appetitive or aversive associations. The calibration asymmetry, however, was not significantly correlated with total accuracy or total calibration, indicating it was not distorted by a ceiling effect. Consequently, it is better able to measure differences between appetitive and aversive learning for both high and low accuracy participants.

In pp. 29, the ANOVA showed a main effect of Condition, indicating a difference between the experimental and control tasks. This was present in both the total accuracy and total calibration variables. Paired t-tests showed that participants scored higher on the experimental task, $M_{acc} = .68$, $SD_{acc} = .13$, $M_{cal} = 29.06$, $SD_{cal} = 21.08$, than the control task, $M_{acc} = .64$, $SD_{acc} = .13$, $M_{cal} = 21.73$, $SD_{cal} = 19.9$. This was true for total accuracy, $t(143) = 3.81$, $p < .001$, and total calibration, $t(143) = 4.31$, $p < .001$. Participants were therefore better able to accept Gain trials and avoid Loss trials than they were able to categorize Seeds trials and Spores trials. This may be attributed to the experimental task being more motivating than the control task, due to its additional risk of losing points. Regardless, this difference in task performance does not cause complications for our main hypothesis.

Comparing the Seeds and Spores trials also revealed a significant difference in their total calibration, but not in their total accuracy. Accuracy on Seeds trials, $M = .63$, $SD = .15$, did not differ significantly from accuracy on Spores trials, $M = .65$, $SD = .14$, $t(143) = 1.623$, $p = .107$. Calibration on Seeds trials, $M = 18.60$, $SD = 23.02$, was significantly lower than that of Spores trials, $M = 24.60$, $SD = 21.98$, $t(143) = 3.397$, $p = .001$. This suggests that, while participants were not able to identify Seeds trials better than Spores trials, the Spores trials that they did identify, they had higher confidence in. Given that the stimuli used for Seeds and Spores were counterbalanced, this cannot be explained by one set of stimuli being more salient than the other. Rather, it appears that there is a slight difference in salience between the terms themselves, in spite of being very similar. This also does not cause complications for our main hypothesis, as even with an intrinsic difference in calibration, none of the predictors used in the regression analyses explained a significant amount of variance in the calibration asymmetry for the control task.

**Total accuracy and calibration**

Supplementary Table 3 shows the bivariate correlations for the variables used in the regression analysis, plus the total accuracy and calibration. There is one result of note: Neuroticism shows a significant correlation with total calibration on the experimental task, and with total accuracy and calibration on the control task. There is no corresponding correlation in the asymmetry variables, indicating that participants high in Neuroticism performed overall worse on the task, but not specifically worse on Gain or Loss trials. This may be ascribed to a number of influences (e.g. attention, motivation, memory) that we are unable to further elucidate, and as it is outside the scope of our main hypothesis, this effect was not analyzed further.

**Performance over consecutive blocks**

To investigate the added effect of impulsivity, a mixed model regression analysis was performed using the calibration asymmetry on each block as a dependent variable and using block and impulsivity as predictors. Several models were compared, and the model using a fixed intercept of block and a random intercept of Impulsivity * Block was shown to have the best fit. Univariate analysis on this model revealed a significant main effect of Impulsivity, $F(1,136) = 8.53$, $p = .004$, but no significant main effect of Block, $F(1,136) = 1.16$, $p = .28$, and no significant interaction effect, $F(1,136) = 1.99$, $p = .16$. As before, the main effect of impulsivity reflected more positive calibration asymmetry values for higher impulsivity scores.

Although this model suggests that there was no change in the magnitude of the calibration asymmetry across blocks, due to the changes in block 3 shown by Figure 3, these differences were investigated further. Three levels for impulsivity were defined based on the mean, $M = 10.83$, and the values two standard deviations, $SD = 1.95$, above and below the mean, allowing comparisons of high, average, and low levels of impulsivity. Contrasts between these levels at each block indicated significant differences at block 1, $t(385) = -3.24$, $p = .004$, and block 2, $t(385) = -3.31$, $p = .003$, but not at block 3, $t(385) = -1.18$, $p = .47$. Estimated marginal means indicated that this was the result of both high and average impulsivity scorers showing a reduction in calibration asymmetry, while low impulsivity scorers showed a minor increase, causing the effect of impulsivity on calibration asymmetry in block 3 to be reduced to non-significance (see Figure 4).

**Supplementary Table 3**

*Bivariate Correlations for Study Variables*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Experimental task | | | | | | | | | | |
| 1. Total accuracy | — | | | | | | | | | |
| 2. Accuracy asymmetry | -.152 | — | | | | | | | | |
| 3. Total calibration | .952** | -.128 | — | | | | | | | |
| 4. Calibration asymmetry | -.011 | .845** | -.023 | — | | | | | | |
| Control task | | | | | | | | | | |
| 5. Total accuracy | .468** | -.051 | .468** | -.023 | — | | | | | |
| 6. Accuracy asymmetry | .004 | .171* | .008 | .200* | .016 | — | | | | |
| 7. Total calibration | .467** | -.045 | .504** | -.039 | .930** | .014 | — | | | |
| 8. Calibration asymmetry | .033 | .127 | .028 | .144 | .062 | .810** | .060 | — | | |
| Predictors | | | | | | | | | | |
| 9. Neuroticism | -.151 | -.001 | -.190* | .011 | -.188* | .010 | -.217** | -.066 | — | |
| 10. Impulsivity | -.151 | .376** | -.136 | .315** | -.058 | -.152 | -.028 | -.173* | .070 | — |
| 11. Anhedonia | .094 | .054 | .107 | -.007 | .108 | -.172* | .119 | -.080 | -.149 | .138 |

Note: $^{*}p < .05$. $^{**}p < .01$.

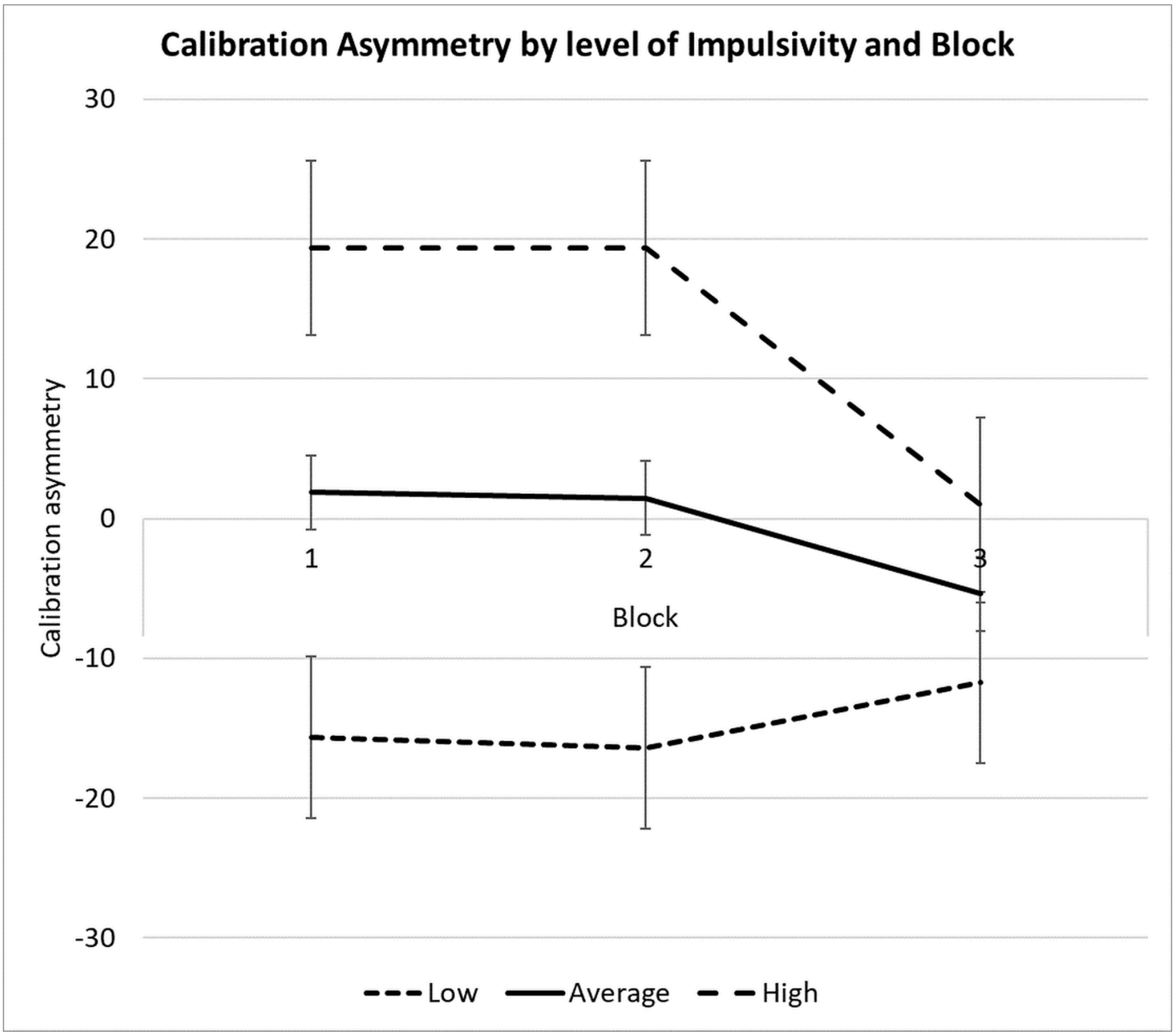



*Figure 4: Estimated marginal means of calibration asymmetry over blocks 1 to 3, split by low, average, and high impulsivity scores. A more positive asymmetry indicates better learning of appetitive than aversive associations, while a more negative asymmetry indicates the reverse.*

These results showed that participants' performance over time showed the strongest increase after the third reinforcement, which also corresponded to significantly better learning of aversive associations. Furthermore, these changes were not equal across different degrees of impulsivity. This is consistent with the finding that, by the end, participants learned aversive associations better on average: as the effect of impulsivity was specific to aversive associations, the differential effect disappeared. Taken together, these results suggest that learning effects became sufficiently strong to supersede the influence of impulsivity. This confirms that the use of ‘weak’ associations is an effective method for finding individual differences, as argued by Beckers et al., (2013) and Lissek et al., (2006), and indicates that learning effects at different levels of reinforcement are relevant to detecting any effects of personality. Thus, it is crucial to design conditioning experiments which aim to detect such effects with care.

# Chapter 3

Distress is Not Delicious: Appetitive Conditioning is Weaker with High Psychological Distress

Published as: Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. (2025). Distress is not delicious: Appetitive conditioning is weaker with high psychological distress. *Journal of Experimental Psychopathology*, *16*(1), 20438087251314526. https://doi.org/10.1177/20438087251314526

**Abstract**

Appetitive and aversive conditioning seemingly play a role in the development and maintenance of various psychopathologies, including anxiety, mood, eating, and substance use disorders. However, studies on conditioning typically only study either appetitive or aversive conditioning in the context of psychopathology, and they are poorly integrated. In this study, 80 healthy volunteers performed both an appetitive and an aversive conditioning task, in which they associated complex 3D objects with appetitive or aversive tastes. An individual measure of learning asymmetry was calculated by comparing their expectancy ratings for these tastes, which was examined in relation to self-report scales on psychological distress, substance use frequency, impulsivity and anhedonia to determine whether stronger learning asymmetry is associated with more symptoms of psychopathology. It was found that learning asymmetry was significantly associated with psychological distress ($R^2$ = .05). Aversive learning showed no difference related to distress, but weaker appetitive learning was associated with higher distress. Substance use, impulsivity and anhedonia showed no relation to learning asymmetry. These findings suggest that relative differences in appetitive and aversive learning may play a role in the sensitivity to psychopathology.

## Introduction

Throughout human history, the skills required to thrive in life have been in constant flux, yet certain cognitive processes have always remained indispensable. One of these is the ability to learn from and respond to reward and punishment, which includes adaptive behaviors that remain essential to navigate in the modern world. However, learned behaviors can be both adaptive and maladaptive depending on the circumstances. Some individuals may seek out appetitive outcomes more than they avoid aversive outcomes, which may play a role in the development of addictive behaviors including gambling (Albrecht et al., 2007) and overeating (Dingemans et al., 2009; Meyer et al., 2015). Others may be oversensitive to aversive outcomes and fail to take advantage of opportunities for appetitive outcomes, which may be associated with excessive fears (Briere et al., 2010). As such, the study of appetitive and aversive learning is important to the understanding of mental health.

Appetitive and aversive learning is frequently studied in the laboratory. Core to this research are conditioning paradigms, which usually include repeated pairings of neutral stimuli (the conditioned stimulus or CS) together with rewarding, appetitive stimuli or unpleasant, aversive stimuli (the unconditioned stimulus or US). Learning is then measured by a change in behavioral and/or physiological responses to the CS. Conditioning studies using aversive stimuli causing fear or pain have long been used to determine whether these conditioned responses are implicated in the development or maintenance of fear and anxiety disorders through, for example, higher responsivity to aversive stimuli (Pitman & Orr, 1986) or persisting avoidance after extinction (Eysenck, 1979). Evidence from different meta-analyses shows that anxiety disorder patients have increased fear responses to the CS- compared to healthy controls, suggesting a reduced discrimination ability for safety cues (Duits et al., 2015; Lissek et al., 2005). Similarly, Pittig et al. (2018) reviews studies that show that anxiety disorder patients respond more strongly to neutral stimuli after having been exposed to aversive stimuli, although others fail to replicate this result, and notes that investigating individual differences related to associative learning is a critical next step towards understanding anxiety disorder. This shows that much ambiguity remains about the role of neutral and aversive conditioned stimuli in anxiety, which is also true for studies focusing on mental disorders other than anxiety: discrimination between aversive and neutral stimuli is shown to be weaker in men at risk for alcoholism (Finn et al., 1994) but stronger in flying phobia patients (Vriends et al., 2012). Such studies are scarce compared to those on anxiety disorder, and more research on the relevance of aversive conditioning to other disorders would be helpful in clarifying how differences in learning behavior can affect psychological functioning.

In contrast to aversive conditioning, there is comparatively little research on how disordered appetitive conditioning may be related to psychopathology. A review has shown that depression and schizophrenia are both associated with impaired appetitive conditioning (Martin-Soelch et al., 2007), which suggests that the development of these disorders may be affected by differences in appetitive learning. Other studies have investigated eating disorders and obesity, with some showing that overweight/obese individuals show poorer discrimination of appetitive (CS+) and neutral (CS-) stimuli than those with healthy weight (Coppin et al., 2014; van den Akker et al., 2017; Zhang et al., 2014), and others finding that appetitive learning is stronger in overweight than in lean individuals (Meemken et al., 2018; Meyer et al., 2015), yet others report no learning differences depending on weight (van den Akker et al., 2019). Given that different studies report inconsistent results regarding discrimination learning of appetitive and aversive stimuli, it may be that the relative values of these stimuli are relevant to an individual's responses. In other words, what is perceived as an appetitive or an aversive stimulus in a given experiment may depend on individual differences: in studies using a CS-, the absence of an appetitive stimulus may be experienced as aversive, or vice-versa for an aversive stimulus. Therefore, this may be clarified by studying the strength of both appetitive and aversive conditioning simultaneously.

When it comes to the study of conditioning and its relevance to psychopathology, most studies are limited to measures of either appetitive or aversive conditioning, with only a few studies that have investigated both types of learning simultaneously. Moreover, when both are measured in one sample, psychopathology is typically not considered, focusing rather on fundamental aspects of appetitive and aversive conditioning (Andreatta & Pauli, 2015; Baeyens et al., 1990; Gera et al., 2019; Kerkhof et al., 2011). The study by Zbozinek et al. (2021) forms a rare exception to this. Here, the authors used Pavlovian conditioning with electric shocks and monetary rewards using an occasion setting paradigm, in which the presence or absence of a sound cue would predict whether a particular CS would be paired with a US. It was found that high trait anxiety was associated with higher fear self-report, US expectancy, and SCR on CS+ trials with electric shocks, and that high trait depression was associated with lower US expectancy on CS+ trials with monetary reward. This leaves an open question of whether these traits are associated with a relative rather than absolute difference in appetitive and aversive learning, which may have more explanatory power for their role in depression and anxiety.

In short, investigating appetitive and aversive learning simultaneously may clarify inconsistencies surrounding the findings on conditioning and psychopathology. In a previous study (Kemp et al., 2024), this was studied in relation to personality traits tied to certain psychopathological characteristics. A conditioning task was used in which participants responded to stimuli based on whether they would gain or lose points, which were tied to an additional monetary reward. Based on these responses, the difference in performance between gain and loss trials was calculated, which represents a learning asymmetry, a measure of participants' bias towards learning from appetitive versus aversive associations. Results showed that learning asymmetry was significantly associated with a self-report measure of impulsivity, but not with neuroticism or anhedonia. It was shown that more impulsive individuals performed worse than less impulsive individuals in learning from punishment trials but they did not learn better from reward trials. These data suggest that the relative strength of appetitive vs. aversive associations may be relevant to the study of approach and avoidance-related behaviors and impulse control disorders. As such, these results warrant further investigation as to whether different types of psychopathologies are consistently associated with a learning asymmetry in either direction.

The current study aims to extend the previous results by investigating whether learning asymmetry is associated with psychopathological symptoms. To this end, a variety of self-report measures was used. First, the Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983) was included as a measure of the presence and intensity of various mental disorder symptoms. Second, the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST; WHO, 2002) was used to measure the frequency and severity of substance use. Third, the UPPS-P Impulsive Behavior Scale (Lynam et al., 2006) was used to assess the degree of impulsive personality traits. Finally, the Snaith-Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995) was used to measure the presence of anhedonia. It was hypothesized that individuals who show a stronger learning asymmetry will show a higher degree of psychopathology, although we have no expectation about whether this will result from a positive or negative learning asymmetry effect. In addition, concerning the personality scales, it is hypothesized that higher levels of impulsivity are associated with a positive learning asymmetry, and higher levels of anhedonia are associated with a negative learning asymmetry. Such effects can be further investigated by determining whether appetitive or aversive learning is stronger or weaker, either or both of which could drive differences in learning asymmetry. Consequently, we expect high impulsivity to be associated with weaker aversive learning, as described in Kemp et al. (2024), and participants high in anhedonia are expected to show decreased sensitivity to appetitive stimuli.

In contrast to our previous study (Kemp et al., 2024a), the current study uses primary reinforcers (i.e., biologically salient stimuli) rather than secondary reinforcers (e.g., money) because primary reinforcers have been shown to result in stronger conditioned responses than secondary reinforcers (M. Delgado et al., 2011). While some previous studies have compared appetitive and aversive primary reinforcers directly (e.g. by administering food rewards or electric shocks; Andreatta & Pauli, 2015), this may pose a problem if the difference in responses is to be quantified: the appetitiveness of food may be qualitatively different from the aversiveness of pain, meaning that comparisons between individuals suffer also from varying individual sensitivity to these different domains (see also van der Schaaf et al., 2022). For this reason, the unconditioned stimuli in this study were both in the domain of taste, namely sweet and bitter-tasting liquids, calibrated to the participants' taste sensitivity. In this way, the learning asymmetry could be determined using the association strength of appetitive and aversive stimuli in the same domain.

## Method

### Participants

In total, 80 undergraduate students participated in the study. They were recruited through the university's research participation system for psychology students. Participants were between 17 and 30 years of age ($M$ = 20.69, $SD$ = 2.91) and consisted of 15 men and 65 women. Participation occurred between 18 November 2022 and 16 February 2023. Participants were compensated with either study credits or a €10 voucher.

### Materials

#### *Stimuli and apparatus*

Three variants of appetitive stimuli were created by diluting concentrated fruit syrup in orange, strawberry, and raspberry flavors (producer: Karvan Cevitam/Kraft Heinz, Chicago, United States) with tap water in a ratio of one part syrup to seven parts water. Five variants of aversive stimuli were created by dissolving powdered quinine hydrochloride (producer: Arnold Suhr, Hilversum, the Netherlands) in tap water to create 0.1, 0.2, 0.3, 0.4, and 0.5 mM solutions. These solutions were administered using a custom-built peristaltic pump, which delivered 0.8 ml of liquid per reinforced trial. To determine which of the appetitive and aversive stimuli should be selected for the participant, samples of 3 ml were presented to the participant, which they were instructed to rate in the order of strawberry-orange-raspberry-0.1 mM-0.2 mM-0.3 mM-0.4 mM-0.5 mM. Each sample was rated on two 7-point scales, one for pleasantness which ranged from 'Extremely unpleasant' (1) to 'Extremely pleasant' (7), and one for intensity, which ranged from 'Tasteless' (1) to 'Extremely intense' (7). Pilot testing ($N$ = 14) indicated that both the appetitive and aversive stimuli were rated as more unpleasant after the task than before. Therefore, the unconditioned stimuli that were chosen for each participant were a) the appetitive stimulus that was rated as the most pleasant and most intense, and b) the aversive stimulus that was rated as most unpleasant (but no higher than 'Moderately unpleasant' (2)) and most intense. In case of ties among these ratings, the participant was asked which appetitive stimulus they preferred, and the highest concentration of the aversive stimulus that received identical ratings was used.

Conditioned stimuli consisted of forty different images of 3D objects generated according to the method of Watson et al. (2019). These objects were made to be abstract and complex, and each object used a combination of variants in shape, color, and surface texture (see Figure 1). A large variety of conditioned stimuli was used to make it sufficiently difficult for participants to learn each association so that individual differences could be best observed, or in the words of Lissek et al. (2006), to create a 'weak' situation. These were divided into two sets of twenty objects. All objects in a set were presented in a randomized sequence. In one set, ten of the objects were paired with the aversive US, and the other ten were paired with no US. For the second set, ten of the objects were paired with the appetitive US, and the other ten with no US. A 2x2 counterbalancing scheme determined whether any object was paired with a US or with no US and whether it was part of the appetitive or aversive set.

***Expectancy ratings***

After a CS was presented at the start of a trial, the participant was instructed to rate their expectancy on whether an unconditioned stimulus would be administered by selecting a point on a 100-point visual analog scale (VAS), which consisted of a horizontal black line 500 by 10 pixels in size. The question "How much do you expect the taste?" was displayed below the CS. The left extreme of the VAS was labeled "Certainly not" (0) and the right extreme was labeled "Certainly" (100). After the participant clicked a point on the VAS, a red line indicated their response.

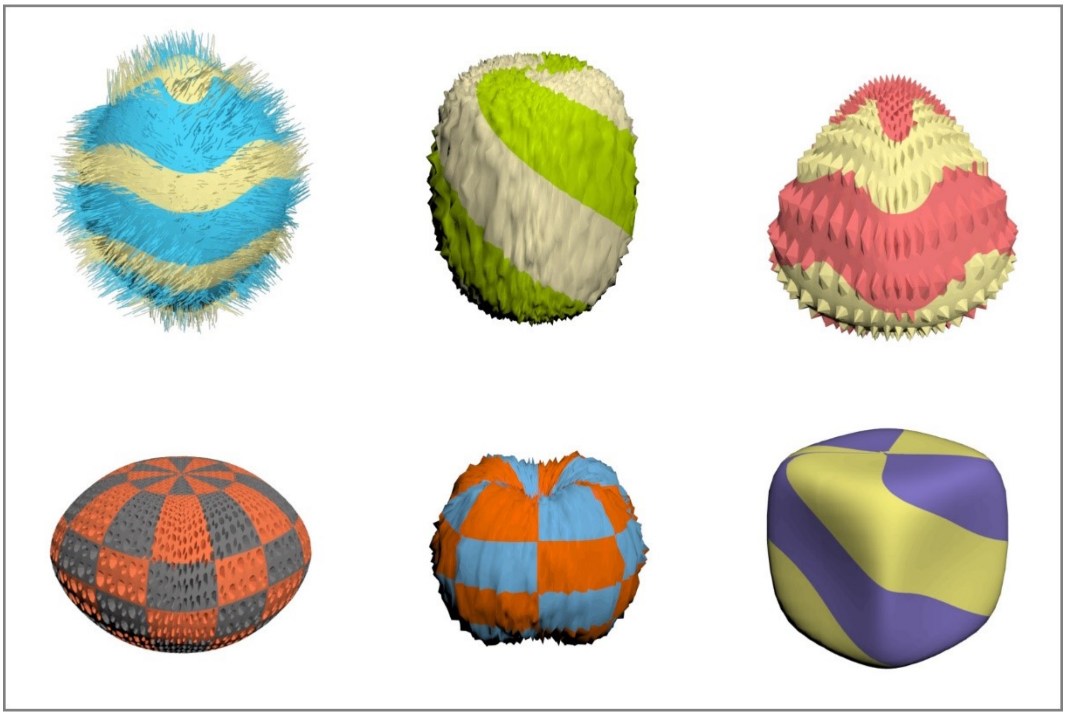


*Figure 1: A selection of six objects out of the forty used for conditioned stimuli.*

***Brief Symptom Inventory (BSI)***

The BSI is a self-report scale for psychological distress. It consists of 53 items rated on a 5-point scale, assessing how much the participant was distressed by a given symptom during the past seven days, ranging from 0 (Not at all) to 4 (Extremely). The items cover nine symptom dimensions: Somatization, Obsession-Compulsion, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. The Global Severity Index (GSI) is calculated using the mean of all 53 items. Hereafter, the GSI is meant when referring to the BSI. Cronbach's alpha for this scale was very high ($\alpha = .959$).

***UPPS-P Impulsive Behavior Scale***

The UPPS-P Impulsive Behavior Scale is an inventory of various kinds of impulsive behaviors, modeled after five factors that were assessed by exploratory and confirmatory factor analysis (Lynam et al., 2006) and consisting of 59 items scored on a scale of 1 (Agree Strongly) to 4 (Disagree Strongly). The subscales include (Negative) Urgency, (lack of) Premeditation, (lack of) Perseverance, Sensation Seeking, and Positive Urgency. The original UPPS included only negative urgency, that is, a tendency to commit rash or regrettable actions as a result of intense negative affect (Whiteside & Lynam, 2001), but its positive affect equivalent was specified later (Cyders & Smith, 2007). While the UPPS-P is commonly analyzed by using each subscale as a factor, to avoid issues of multicollinearity, all subscales were averaged to obtain a total impulsivity score. Cronbach's alpha for this scale was very high ($\alpha$ = .911).

### *Snaith-Hamilton Pleasure Scale (SHAPS)*

The SHAPS is a 14-item self-report scale measuring the reduction or absence of pleasurable experiences, or anhedonia (Snaith et al., 1995). Items are scored on a 4-point scale from 1 (Agree strongly) to 4 (Disagree strongly). All items are summed to obtain a total score. Cronbach's alpha for this scale was adequate ($\alpha$ = .756).

### *The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)*

The ASSIST is a substance abuse screening test that measures the frequency of recreational use of various substances, as well as the degree to which this use results in problems for the participant. It consists of a list of ten categories of substances, and the participant is asked if they ever used this substance recreationally. For each category that they answered in the affirmative, they are then asked six questions about their use of this substance and its impact on their life (e.g. "How often have you used cannabis?"), which are answered in the form of a five-point scale ranging from "Never" to "Daily or almost daily", or on a three point scale with the choices "No, never", "Yes, in the past 3 months", and "Yes, but not in the past 3 months". More frequent use and more risky substances are reflected in higher scores. Cronbach's alpha for this scale was high ($\alpha$ = .887).

## Procedure

The experimental procedure was approved by the local Ethics Review Committee Psychology and Neuroscience. Prior to arriving at the lab, the participant was asked to refrain from eating or drinking anything in the two hours prior to the experiment. Upon arriving, the participant was instructed to read the information letter after which the informed consent was signed. Then the participant performed a taste test to calibrate the stimuli to their taste sensitivity. Following the taste test, the participant was seated in front of a computer screen and provided with verbal instructions on the task. After receiving the instructions, they were provided with a head mount to which the end of a length of silicone tubing (4 mm outer diameter, producer: Saint Gobain, La Défense, France) was affixed. The participant was instructed to hold the tube loosely in their mouth. Then, they were supervised by the experimenter while they performed twelve practice trials using six 3D objects, presented in the sequence 1-2-3-4-5-6-1-2-3-4-5-6, in which 1 to 3 were reinforced, and 4 to 6 were unreinforced. The 3D objects used for the practice trials were not presented in the actual task. During the practice trials, the participant received whichever US would also be given in the first block of the task, as determined by the counterbalancing scheme. After the practice trials, the participant performed the conditioning task, which consisted of four blocks with the appetitive stimulus and four blocks with the aversive stimulus. Due to a strong aftertaste induced by the quinine solution, the appetitive and aversive stimuli were each presented in their own phase of the experiment. Half of the participants received the appetitive blocks first, and half received the aversive blocks first.

Each block consisted of ten reinforced (CS+) trials, in which 0.8 ml liquid (US) was administered per trial, and ten unreinforced (CS-) trials, in which no liquid was administered (no US). On each trial, the participant was presented with a CS, and then rated their expectancy for whether a US would follow. After their response, they received either the US or no US, during which the CS remained visible. This was repeated for twenty trials, with each trial presenting a different CS. After twenty trials, the participant was allowed to pause and choose when to continue onto the next block of trials and have a drink of water if desired. In the following two blocks, the same twenty CSs were repeated, each again followed by the US or no US. On the fourth and final block, each CS was again presented as before, but the US was omitted from all trials. Participants were instructed that they would not receive any US in this block and to respond with their expectancy as in previous blocks. This was done to prevent the USs, particularly the bitter taste, from becoming too unpleasant over time. Given that no CSs were repeated within blocks, the presentation of the US did not provide participants with information on the other CS-US associations in that block, and therefore the omission of the US in the fourth block was not expected to affect participants' responses. The presentation order of all CSs was randomized between participants and slightly shuffled between blocks, so that each CS occupied roughly, but not exactly, the same position in each block.

After the participants completed the task, they filled out the self-report measures, including questions about their effort, understanding, and experience of the task. Details on these questions are available in Supplementary Table 1. These were followed by demographics questions, the SHAPS, the UPPS-P, the BSI, and finally the ASSIST. Then, the participants received their compensation and were thanked for their participation.

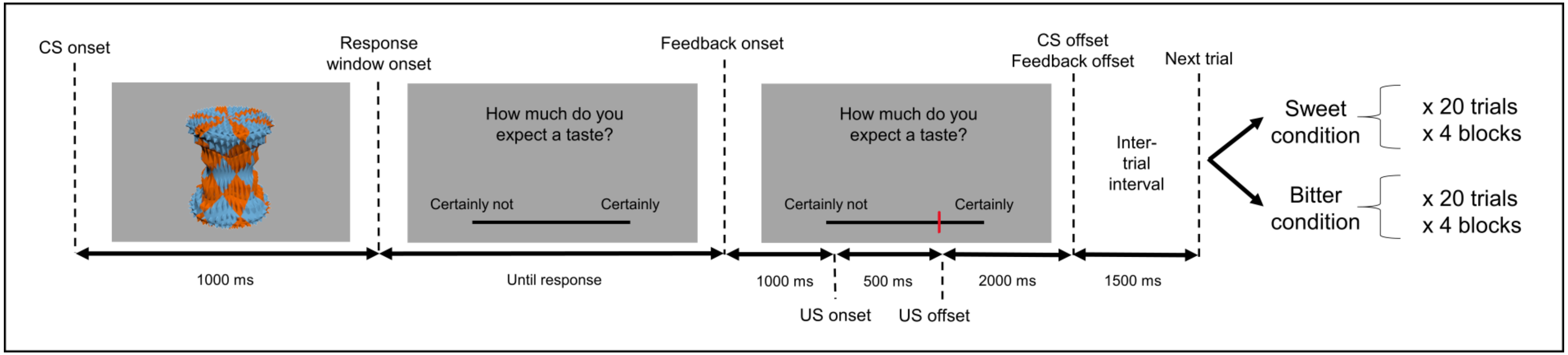



*Figure 2: Diagram of the task procedure. Areas in grey represent elements displayed during the task, simplified for readability. Feedback was given as a red line on the point on the VAS where the participant clicked.*

**Statistical analyses**

All analyses were performed in IBM SPSS Statistics version 26, except for those in the sections 'Development of the learning asymmetry across blocks' and 'Estimated Marginal Means of BSI: Learning Asymmetry per Block', which were performed in R Studio, using the packages lme4 and emmeans. Graphs were created in R Studio using the packages ggplot2, ggExtra, patchwork and gridExtra. Participants' ability to distinguish the CS+ from the CS- was determined by first separating trials by valence (appetitive/sweet or aversive/bitter) and then subtracting all expectancy ratings of CS- trials from CS+ trials and dividing this by the number of trials. Then, their learning asymmetry was determined by subtracting the value obtained from the aversive learning phase from the value obtained from the appetitive learning phase. This creates a measure that is more positive if the participant has a higher expectancy for the CS+ relative to the CS- on appetitive trials, and more negative if the same is true for aversive trials. Trials from the second, third and fourth blocks of the task were used for this measure, as the first block was used as an acquisition phase and participants responded without having had the opportunity to learn.

To determine the relative contribution of each variable to the severity of mental health symptoms, two hierarchical regression analyses were performed using the learning asymmetry as the dependent variable, the BSI and the ASSIST as predictors in the first regression analysis, and the UPPS-P and the SHAPS as predictors in the second regression analysis. Two separate regression analyses were used to simplify the interpretation of the results, as shared variance between personality measures and psychological distress measures may create ambiguity about their relative contribution to learning asymmetry. Then, to determine whether and how any of these effects changed throughout each learning phase, a mixed model regression analysis was performed using the learning asymmetry against any predictor variables that showed a significant effect in the previously performed regression analyses.

## Transparency and openness

The study's hypotheses, analyses, exclusion criteria, and sample size were preregistered at https://aspredicted.org/f35s-8tgd.pdf. All data, analysis code, stimulus materials, and task script are available at https://osf.io/cqsgm/. Using G*Power version 3.1.9.7 (Faul et al., 2007) for an *a priori* power analysis, it was determined that to achieve a power of 0.80 for an effect size of 0.13 (based on Kemp et al., 2024) at a significance criterion of $\alpha = .05$, a minimum sample size of 78 would be required for a linear multiple regression analysis with two predictors. Exclusion criteria were specified as follows: the appetitive and aversive US not being rated as pleasant or unpleasant respectively, or both tastes not being rated different from each other. Learning asymmetry and self-report values that fell more than 2.5 *SD* outside the mean were winsorized.

# Results

## Inclusion and Exclusion

Based on the average learning asymmetry across all blocks, one participant had their value recoded to the closest non-outlier value, and based on the learning asymmetries per block, three participants had their values recoded. No participants met the exclusion criteria.

## Task Performance

To verify that participants learned which CS was associated with which taste, a repeated-measures analysis of variance (ANOVA) was performed on each participant's averaged expectancy ratings using Block (1-4), Valence (Aversive, Appetitive) and Trial (CS+, CS-) as within-subjects factors. The assumption of sphericity was violated for some tests, and so the Greenhouse-Geisser correction was applied. Test statistics can be found in Table 1.

**Table 1**

ANOVA test statistics for main and interaction effects of expectancy ratings

| Effect | *F* | *df* | *p* | $\eta_p^2$ |
|---|---|---|---|---|
| Valence | 0.92 | 1, 79 | .340 | .012 |
| Trial | 222.18 | 1, 79 | <.001 | .738 |
| Block | 11.09 | 3, 77 | <.001 | .302 |
| Valence * Trial | 0.73 | 1, 79 | .396 | .009 |
| Valence * Block | 0.07 | 3, 77 | .964 | .003 |
| Trial * Block | 48.50 | 3, 77 | <.001 | .654 |
| Valence * Trial * Block | 0.003 | 3, 77 | .999 | .000 |

Significant main effects were found for Trial and Block, in addition to a significant interaction effect between Trial and Block. Comparisons of the means showed that this interaction corresponded to expectancy ratings for CS+ trials increasing with every block, and those for CS- trials decreasing with every block. Using paired samples *t*-tests for CS+ and CS- expectancy ratings, we confirmed that they were significantly different for all blocks but the first (see Figure 3). The non-significant main effect of Valence indicated that the sample did not show an overall bias towards either appetitive or aversive learning. It was concluded that the sample showed robust learning for both the appetitive and the aversive tastes. Further tests can be found on pp. 73.

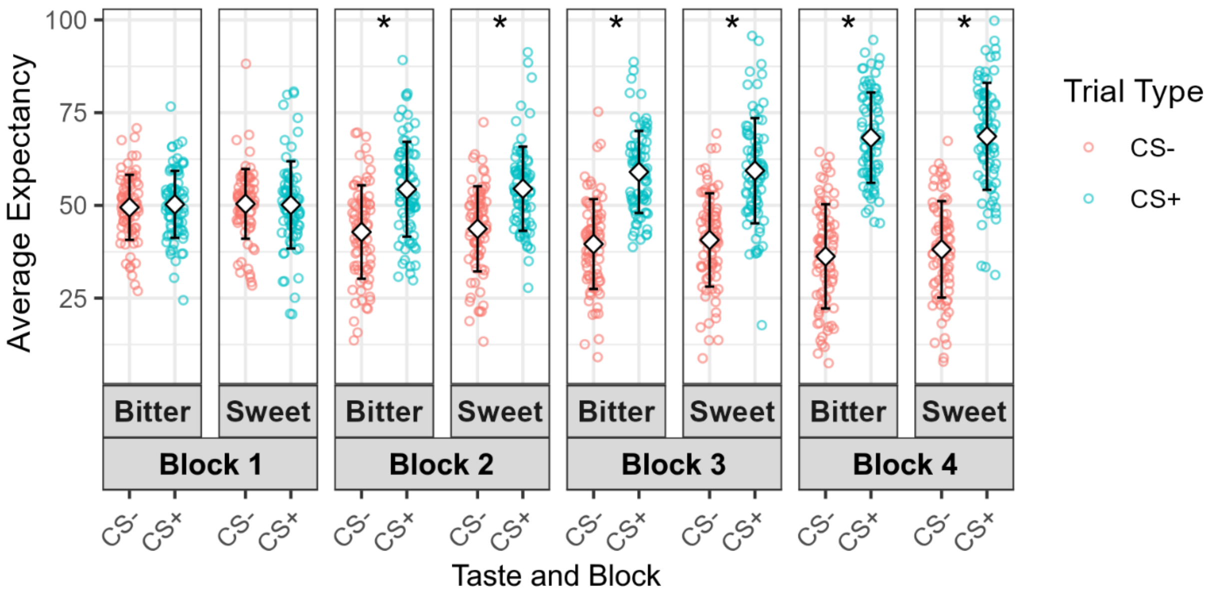



*Figure 3: Comparison of expectancy ratings by taste, block and trial type. Circles indicate participant means. Diamonds indicate sample means. Error bars indicate sample standard deviations. Asterisks indicate significant differences between sample means of CS+ and CS- (p < .05, Bonferroni-corrected).*

**Learning asymmetry and self-report measures**

Descriptive statistics for the self-report measures can be found in Supplementary Table 1. To obtain the learning asymmetry values, expectancy ratings for all CS- trials in one block were subtracted from expectancy ratings for all CS+ trials in one block. The difference scores of each bitter taste block were then subtracted from the difference scores of the corresponding sweet taste block. In a previous study (Kemp et al., 2024a), the effect of impulsivity was only found in block 2 and 3, not 4, indicating that impulsivity no longer affected responses once the associations were sufficiently strong. This indicated that impulsivity-related individual differences in learning were measurable starting in block 2. Given that we analyze learning asymmetry in relation to several self-report measures, the average of blocks 2, 3 and 4 was chosen to ensure the relevant learning effects could be measured. Block 1 was omitted from the average, as it served as the acquisition block. A one-sample t-test showed that this average did not differ significantly from zero, $t(79) = -0.63$, $p = .531$, and a Kolmogorov-Smirnov test showed that the distribution did not significantly deviate from normality, $D(80) = 0.066$, $p > .200$. Furthermore, the analysis on pp. 59 indicated that aversive trials did not differ from appetitive trials overall in the expectancy ratings they received.

Two linear regression analyses were performed using the learning asymmetry as the dependent variable. First, the UPPS-P and the SHAPS were used as predictors, representing impulsivity and anhedonia, respectively. Adding the UPPS-P score to the model resulted in no significant change in the explained variance, $F_{change}(1, 78) = 0.71$, $p = .236$, $R^2 = .018$. Adding the SHAPS score also did not significantly change the model, $F_{change}(1, 77) < .01$, $p = .991$, $R^2 = .018$. The final model including two predictors was not significant, $F(2, 77) = 0.71$, $p = .497$. For the second regression analysis, the BSI and the ASSIST were used as predictors, representing general mental health. Adding the BSI score explained a significant amount of the variance, $F_{change}(1, 78) = 4.08$, $p = .047$, $R^2 = .05$. Adding the ASSIST score, however, did not significantly change the explained variance, $F_{change}(1, 77) = 0.38$, $p = .541$, $R^2 = .054$. The final model including two predictors was not significant, $F(2, 77) = 2.21$, $p = .116$. As a final check, all predictors were entered into one regression using the backwards entry method. The UPPS-P, SHAPS and ASSIST did not meet the criteria for inclusion, leaving only the BSI. Thus, the BSI was significantly associated with the learning asymmetry, showing that higher scores were associated with a more negative learning asymmetry, and the other questionnaires did not show a significant association. Further model statistics can be found in Table 2.

**Table 2**

*Summary of hierarchical regression analysis for self-report measures predicting learning asymmetry*

| Model | Variable | *B* | SE *B* | *β* | SE *β* | *t* | *p* | 95% CI for B | | Tolerance |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper | |
| A | Step 1 | | | | | | | | | |
| | UPPS-P | 8.8 | 7.4 | .13 | .11 | 1.20 | .236 | -5.8 | 23.4 | 1.00 |
| | Step 2 | | | | | | | | | |
| | UPPS-P | 8.8 | 7.4 | .13 | .11 | 1.19 | .239 | -6.0 | 23.5 | 1.00 |
| | SHAPS | 0 | 0.7 | 0 | .11 | -0.01 | .991 | -1.3 | 1.3 | 1.00 |
| B | Step 1 | | | | | | | | | |
| | BSI | -8.5 | 4.2 | -.22 | .11 | -2.02 | .047 | -16.8 | -0.12 | 1.00 |
| | Step 2 | | | | | | | | | |
| | BSI | -7.7 | 4.4 | -.20 | .11 | -1.75 | .084 | -16.4 | 1.0 | .92 |
| | ASSIST | -0.1 | 0.1 | -.07 | .11 | -0.61 | .541 | -0.3 | 0.2 | .92 |
| C | UPPS-P | 8.2 | 7.3 | .13 | .11 | 1.12 | .265 | -6.4 | 22.83 | .99 |
| | SHAPS | 0.3 | 0.7 | .05 | .11 | 0.46 | .650 | -1.04 | 1.66 | .95 |
| | BSI | -7.6 | 4.5 | -.20 | .12 | -1.70 | .094 | -16.47 | 1.32 | .89 |
| | ASSIST | -0.1 | 0.1 | -.09 | .12 | -0.76 | .450 | -0.33 | 0.15 | .90 |

Note: Model A: personality predictors, Model B: psychopathology predictors, Model C: all predictors, CI: Confidence Interval.

To explore the relative components of the learning asymmetry and their association with the BSI, Pearson correlation coefficients (two-tailed) were computed using the difference scores of the aversive trials and those of the appetitive trials, together with the BSI. This showed that the difference score on appetitive trials correlated stronger with the BSI, $r(78) = -.24$, than the difference score on aversive trials, $r(78) = .04$. Steiger's Z-test (Hoerger, 2013) was performed to test the difference between correlations accounting for the correlation between appetitive and aversive difference scores, $r(78) = .14$, and this was marginally significant, $Z_H(77) = 1.92$, $p = 0.055$. While this result alone is not conclusive, the correlation with aversive learning being near zero indicates that appetitive learning plays the larger role in the association between BSI and learning asymmetry. In other words, participants with higher psychological distress showed weaker appetitive learning as compared to participants with lower distress, but did not differ with respect to aversive learning (see Figure 4). Further details can be found on pp. 74-79.

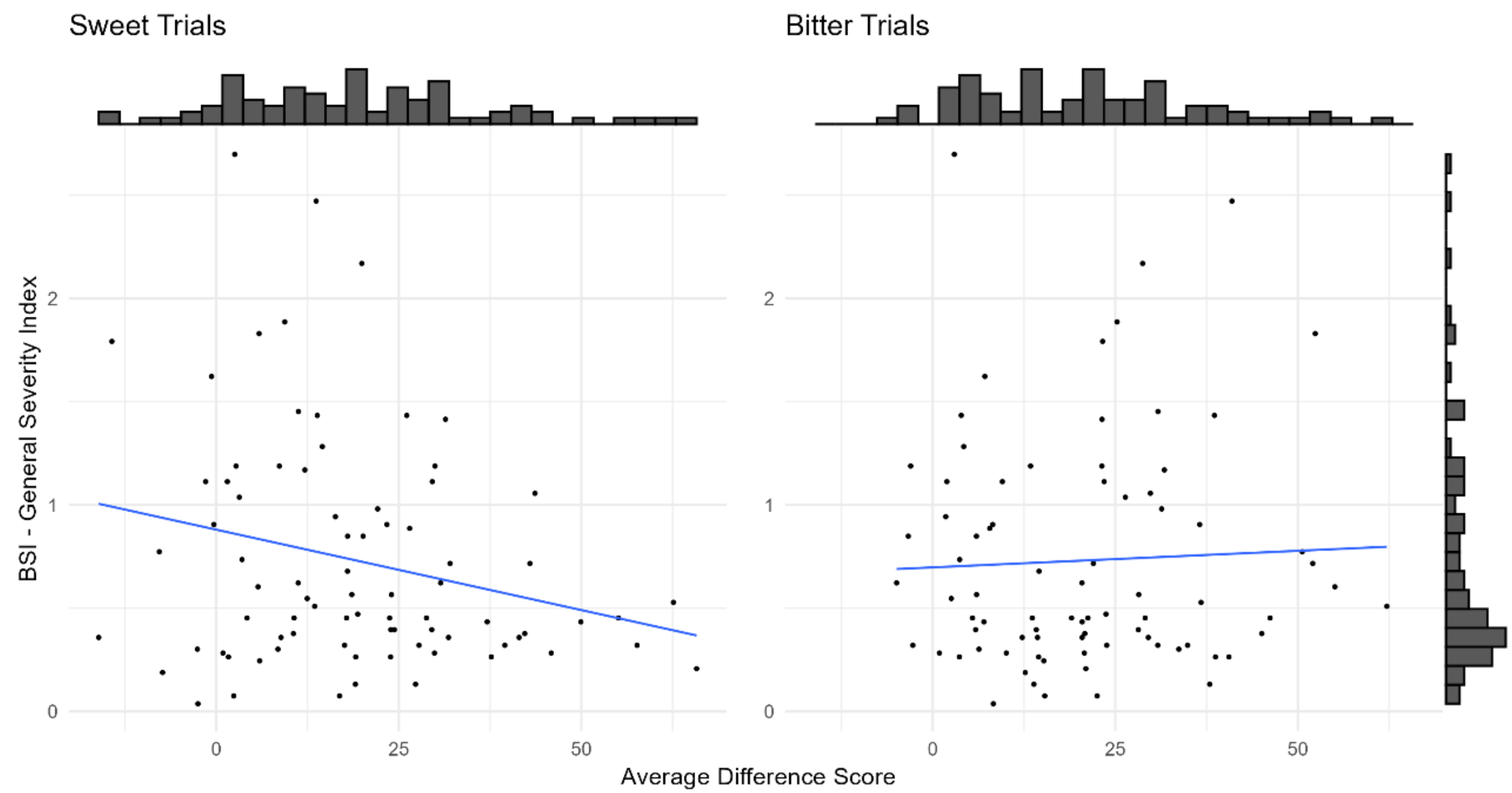



*Figure 4: Graph of BSI scores against average differences between CS+ and CS- trials, separated by sweet and bitter trials. Histograms show marginal distributions. Trend lines show linear models.*

### Development of the learning asymmetry across blocks

To investigate whether the effect of the BSI on learning asymmetry differed by block, a mixed model regression analysis was performed using the learning asymmetry on each block as a dependent variable and using block and BSI as predictors. As the change in learning symmetry across blocks was expected to be similar between participants, and as the variation of BSI across blocks was the effect of interest, these were included as fixed intercepts. Although random intercepts for block were considered, this resulted in overparameterization, and so the final model included fixed effects for BSI, Block, and their interaction BSI x Block. The model also included a random intercept for Participant to account for between-subject variability. Univariate analysis on this model revealed a significant main effect of BSI, $F(1, 78) = 7.00$, $p = .029$, but no significant main effect of Block, $F(3, 234) = 0.32$, $p = .81$, and no significant interaction effect, $F(3, 234) = 0.67$, $p = .57$. As before, the main effect of BSI reflected more negative asymmetry values for higher BSI scores. However, the lack of an interaction effect indicates that there were no significant changes in how strongly the BSI predicted the learning asymmetry over time. In other words, participants with higher BSI scores showed a more negative learning asymmetry even after multiple reinforcements. Further exploratory analysis of per-block effects can be found on pp. 81-82.

## Discussion

In this study, we assessed individual differences in the relative strength of appetitive and aversive learning, and their relation to different measures of personality and mental health. Participants reported their expectancy of whether a CS would be followed by a US or not, and these expectancy ratings were compared to assess the learning asymmetry, a measure of the relative strength of appetitive versus aversive conditioning. Participants showed robust learning effects and considerable variation in learning asymmetry values.

The analysis of the learning asymmetry was twofold: separate regressions were performed on short-term self-reports on psychological distress and substance use, and long-term self-reports on impulsivity and anhedonia. First, psychological distress was found to be associated with a learning asymmetry, but substance use was not. Second, contrary to our hypotheses, neither anhedonia nor impulsivity were found to be associated with learning asymmetry. The significant association between psychological distress and learning asymmetry supports our hypothesis. This effect was driven by a reduction in difference scores for trials in which the participant was administered an appetitive US, indicating that higher psychological distress was associated with weaker appetitive learning, but not with differences in aversive learning. This is consistent with the findings of Shook et al. (2007), who used a different method to measure learning asymmetry and found that participants scoring higher on the Cognitive Style Questionnaire and the Beck Depression Inventory showed weaker appetitive but not aversive learning. Furthermore, in a study using electric shocks and monetary rewards as aversive and appetitive stimuli, Zbozinek et al. (2021) found that trait depression was negatively associated with appetitive learning. Other studies have shown that reward processing is impaired in participants with more depressive symptoms (Bakker et al., 2019). This suggests a general pattern of weaker appetitive learning in individuals with high psychopathology and psychological distress. Future research could further investigate these psychological symptoms to determine whether the effect of impaired appetitive learning is specific to depression or anxiety, or whether it is a transdiagnostic factor for a variety of mental disorders. However, substance use did not add to the explained variance in learning asymmetry after accounting for the effect of psychological distress, which is contrary to our expectations. This may be because substance use and psychological distress were significantly correlated, and their shared variance reduced the likelihood that both measures would show a significant effect with learning asymmetry. While this makes it difficult to draw conclusions about the relation between learning asymmetry and substance use, it does suggest that appetitive learning has some transdiagnostic properties.

It was remarkable that impulsivity was not related to learning asymmetry, because we previously found more impulsive participants showing weaker aversive learning (Kemp et al., 2024a). This may be due to the different designs of the tasks. Like Shook et al. (2007), Kemp et al. (2024) used a task in which participants could gain or lose points depending on the accuracy of their responses, meaning that, unlike the current study, participants had control over the outcome. In other words, this design had more in common with instrumental conditioning than the current study, in which the appetitive and aversive US were always administered, a design closer to Pavlovian conditioning. This may have influenced the responses of impulsive participants in particular. Furthermore, the current task also separated the blocks so that the aftertaste of the aversive US would not contaminate the appetitive US, in contrast to Kemp et al. (2024) where appetitive and aversive trials were intermixed. Gray's Reinforcement Sensitivity Theory (RST; Gray, 1975) may provide an explanation. This theory posits that approach behavior and avoidance behavior are governed by two distinct neural systems, each of which may be activated more strongly or weakly relative to the other, and thus influence an individual's responses toward appetitive and aversive stimuli (Carver & White, 1994). Thus, whether one appropriately responds to a potential reward or threat depends on whether these systems are appropriately 'tuned'. These systems are represented by two traits that govern responses to appetitive and aversive stimuli: anxiety and impulsivity.[2] In this framework, anxiety is thought to mediate responses to both appetitive and aversive stimuli, whereas impulsivity governs behavior towards appetitive stimuli (Corr, 2008). Therefore, impulsivity may not have affected learning when only aversive stimuli were presented, as in the design of the current study, even though it did affect learning when appetitive and aversive stimuli were intermixed, as in Kemp et al. (2024). If this is indeed the case, it suggests that the contrast between appetitive and aversive USs is more important in discovering personality-related effects of the learning asymmetry than the choice of US. Future research should determine whether RST can provide further insight into the importance of appetitive and aversive learning for an individual's propensity for mental disorders.

Although the current study, and likewise Kemp et al. (2024), did not find an association between anhedonia and learning asymmetry, we did find that higher psychological distress was associated with weaker appetitive learning. This is in line with the finding by Shook et al. (2007), who showed weaker appetitive learning associated with depression, anxiety and a negative cognitive style. However, this presents a confusing picture. Anhedonia is one of the central symptoms of depression (APA, 2013), and is described as a lack of enjoyment of pleasant experiences. This is

[2] Commonly measured with the Behavioral Inhibition Scale and the Behavioral Activation Scale (BIS/BAS) respectively.

consistent with the aforementioned results showing weaker appetitive learning. However, this effect was unrelated to anhedonia as measured by the SHAPS. Similarly, psychological distress was not correlated with anhedonia. This may point to a lack of sensitivity in the SHAPS, in that they are unable to detect any effect related to depressive symptoms in a non-clinical sample of volunteers. By contrast, the learning asymmetry does appear to be able to detect an effect on appetitive learning, suggesting that behavioral measures may be more suited to measuring vulnerability to depression in non-clinical populations than self-report measures of anhedonia.

We have sought to minimize extraneous influences in obtaining a measure of appetitive versus aversive learning strength, yet factors such as overall learning strength and confidence may still have affected learning asymmetry. Individual differences in how confidently participants respond could have exaggerated their learning asymmetry, but this was found to be equal between appetitive and aversive learning and did not influence the effect of psychological distress (see pp. 80-81). However, participants' confidence was inferred from how close their expectancy ratings were to 0 or 100, and confidence was not assessed separate from expectancy. Other studies have investigated the effect of confidence, or 'metacognitive awareness' (Leganes-Fonteneau et al., 2018) on Pavlovian conditioning by way of self-report after the conclusion of the task. In the current study, however, the CS-US associations were specifically made to be unpredictable even with full contingency awareness. In addition to this, participants were made aware that no USs would be presented in the fourth block. Although not on purpose, this resembles the more typical method of extinction learning, which measures the decay in participants' CR after US presentations cease. While the absence of feedback in this block was a sharp contrast with the preceding blocks, the strength of their learning could not have been affected by this, as participants could only base their expectancy on feedback presented in previous blocks. Nevertheless, it would be valuable for future research to measure and manipulate awareness, confidence and extinction to investigate appetitive and aversive learning.

Given that this study investigated psychological symptoms in a non-clinical sample of volunteers, our finding may not necessarily translate to clinical mental disorder patients, and additional research is needed to study the relevance of learning asymmetry in clinical mental disorders. In addition, our sample being primarily female may have resulted in over- or under-representation of symptoms and traits for this population. We investigated learning asymmetry as a potential transdiagnostic process, for which we measured different psychopathological symptoms which were aggregated as psychological distress. Furthermore, although a previous study found an association between the learning asymmetry and the UPPS-P (Kemp et al., 2024), the current study did not. This scale asks the participant about their experiences during the past seven days. As such, it may not be stable over longer periods of time, and the association with the learning asymmetry may reflect the measurement of changes in state rather than in trait. Additional research is needed to determine whether learning asymmetry is more trait-like than state-like. Finally, it should be noted that the failure to replicate the effect of impulsivity may be related to the use of taste stimuli, in contrast to the task employed by the previous study which used point gain and loss tied to a monetary reward (Kemp et al., 2024). Individual variation in sensitivity to sweet and bitter taste may have played a role in this, in spite of the steps taken to ensure that the stimuli were balanced between participants of different taste sensitivity. While this complicates comparisons between them, the design of the current study is more in line with other conditioning studies.

In conclusion, these results provide some evidence that the relative differences in appetitive and aversive learning are weakly associated with increased psychopathological symptoms. Furthermore, they show that conditioning experiments using ‘weak’ situations (i.e. ambiguous CS-US associations) could detect individual differences in learning that experiments using ‘strong’ situations may miss (Beckers et al., 2013; Lissek et al., 2006). As demonstrated, using CSs with high complexity is a viable way of measuring the individual differences that may be relevant to an individual’s propensity for developing psychopathology. Our results show the potential for this avenue of research to further explore the role of learning asymmetry in psychopathology.

## Supplementary Material

### Participant Task Evaluation

**Supplementary Table 1**

*Questions and Answers from the Task Follow-Up Questionnaire*

| Question | Answers |
|---|---|
| 1. Were the tasks and instructions clear and understandable? | a) Yes, I understood everything perfectly<br>b) Yes, I understood most of it<br>c) No, there were a lot of things I didn't understand<br>d) No, I didn't understand any of it |
| 2. Please elaborate on what you found unclear. | |
| 3. Did you put a good effort into giving accurate answers? | a) Yes, I did the best that I could<br>b) Kind of, I didn't try as hard as I could have<br>c) Not really, I wasn't paying much attention<br>d) No, I was just clicking randomly |
| 4. Did you use any memorization techniques or other methods that helped you respond accurately? If yes, please describe them. If no, you can leave this blank. | |
| 5. Which of these words match your feelings about the task? Please select all that apply. | ▪ Interesting ▪ Boring<br>▪ Frustrating ▪ Difficult<br>▪ Challenging ▪ Rewarding |

| | |
|---|---|
| 6. How distinguishable did you find the different objects? | a) They were all clearly different<br>b) They were somewhat similar but still distinguishable<br>c) They had a lot of similarities and I mixed some of them up<br>d) They were so similar I mixed them up constantly |
| 7. Are there any other thoughts or opinions you had about the tasks? Please list any that you can think of. | |
| 8. What was the last time you ate before the experiment? | a) One hour before the experiment or less<br>b) Two hours before the experiment or less<br>c) More than two hours before the experiment |
| 9. What was the last time you drank something before the experiment? | a) One hour before the experiment or less<br>b) Two hours before the experiment or less<br>c) More than two hours before the experiment |

Note: Empty cells in the 'Answers' column indicate that participants were allowed to type in their own answer.

In the task-related follow-up questions, all participants indicated that the instructions were mostly or fully understandable and 82.5% indicated putting their best efforts in, with the remainder reporting putting in a moderate effort (Table 2). When asked about the distinguishability of the stimuli, 63.75% of participants reported them to be mostly or very distinguishable, which is expected given that the stimuli consisted of many combinations of a few different properties. As such, these answers indicated that the balance between recognizability and difficulty was well-struck.

**Supplementary Table 2**

*Answer frequency for task-related follow-up questions*

| Question | Answer | | | |
|---|---|---|---|---|
| | Yes, very much. | Yes, mostly. | No, not very. | No, not at all. |
| Were the instructions clear and understandable? | 92.5% | 7.5% | 0% | 0% |
| Did you put a good effort into giving accurate answers? | 82.5% | 17.5% | 0% | 0% |
| Did you find the different objects distinguishable? | 15% | 48.75% | 35% | 1.25% |

Note: Response levels represent an approximation of the multiple-choice answer. For the exact phrasing, see Supplementary Table 1.

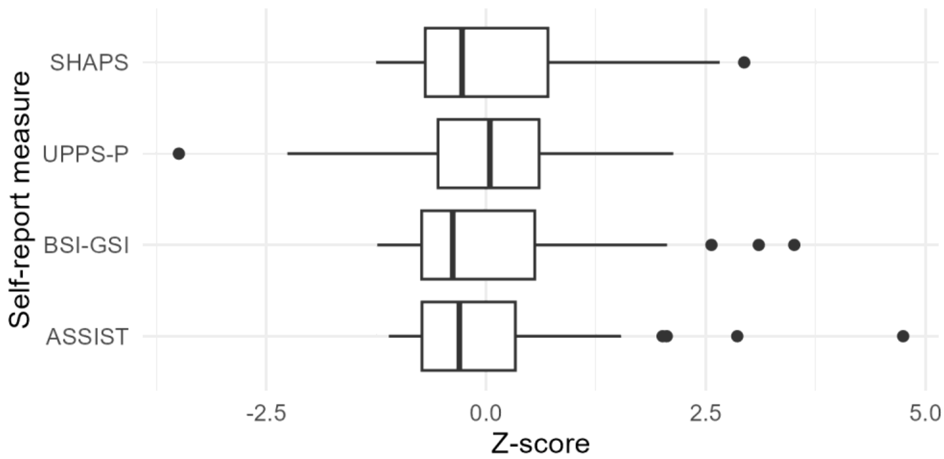
In the answers participants provided to the open questions, nothing indicated a recurrent issue with the task. In reporting when they had eaten prior to the experiment, 63.75% indicated it had been more than two hours, 21.25% indicated it had been between one and two hours, and 8.75% indicated it had been one hour or less. In reporting whether they had drunk anything prior to the experiment, 26.25% indicated it had been more than two hours, 27.5% indicated it had been between one and two hours, and 40% indicated it had been one hour or less. For the eating and drinking questions, 6.25% of participants had missing data. These questions indicated moderate to high compliance to our request to refrain from eating before the experiment, and low compliance to our request to refrain from drinking before the experiment. The latter may have influenced the results, and so Pearson correlation coefficients were calculated between the average difference scores for the aversive block and the appetitive block, and the time since the participant had drunk anything. These correlations were not significant for either the aversive block, $r(75) = .05$, $p = .691$, or the appetitive block, $r(75) = .02$, $p = .838$. From this we conclude that, although we may have been unsuccessful in controlling for the amount that participants drank beforehand, whether they refrained from drinking or not did not influence their performance on the task.

## Self-report Measures

**Supplementary Table 3**

*Descriptive statistics for self-report measures*

| Statistic | BSI | ASSIST | UPPS-P | SHAPS |
|---|---|---|---|---|
| *M* | 0.73 | 23.43 | 2.81 | 18.48 |
| *SD* | 0.56 | 21.19 | 0.33 | 3.58 |
| Range | 2.66 | 124 | 1.83 | 15 |




*Supplementary Figure 1: Box plots showing Z-scores of self-report measures.*

**Supplementary Table 4**

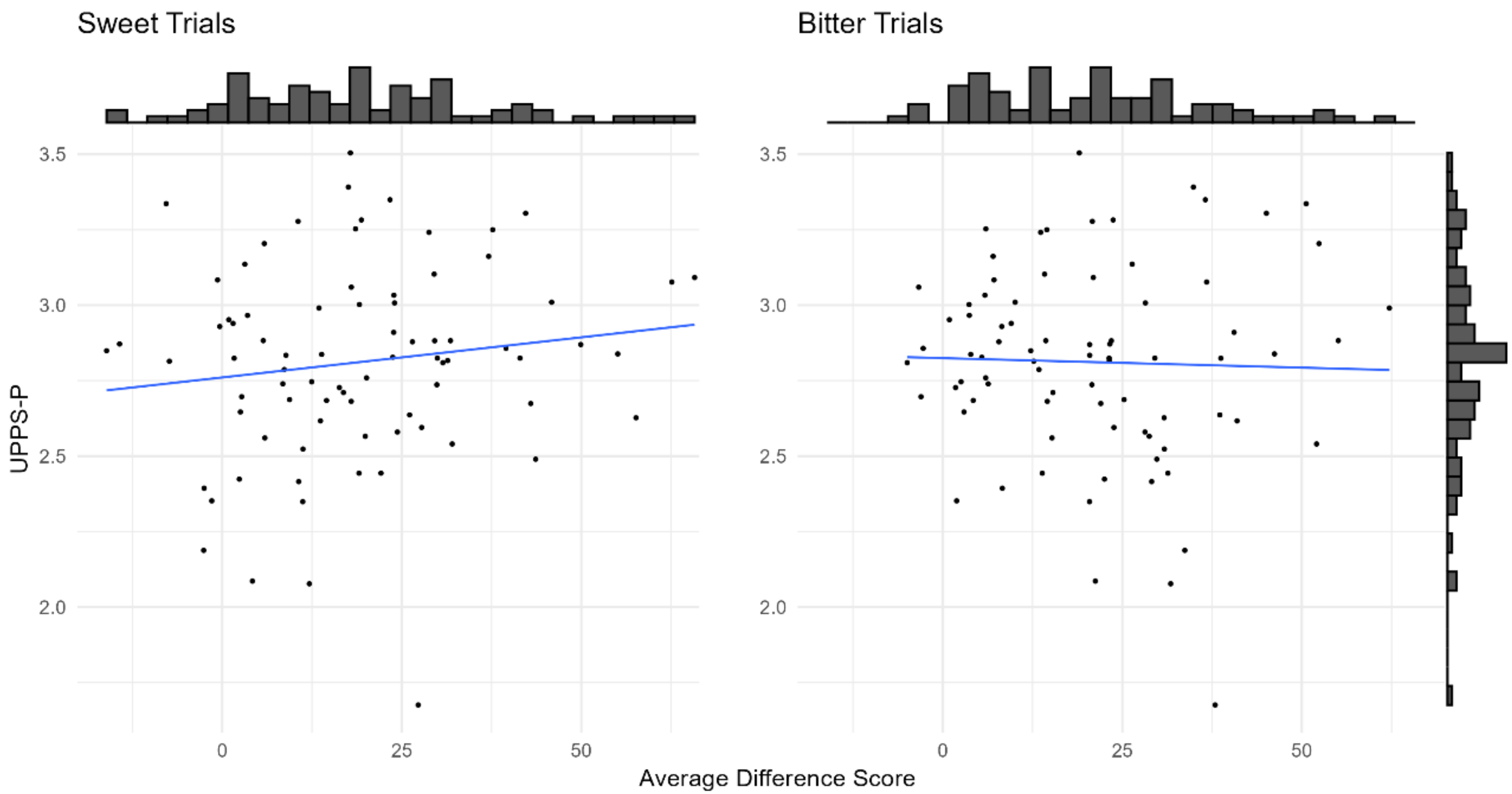
*Bivariate correlations between self-report measures*

| Measure | Learning asymmetry | BSI | ASSIST | UPPS-P | SHAPS |
|---|---|---|---|---|---|
| Learning asymmetry | — | | | | |
| BSI | -.223* | — | | | |
| ASSIST | -.129 | .285* | — | | |
| UPPS-P | .134 | -.061 | .063 | — | |
| SHAPS | .181 | .185 | .181 | .039 | — |

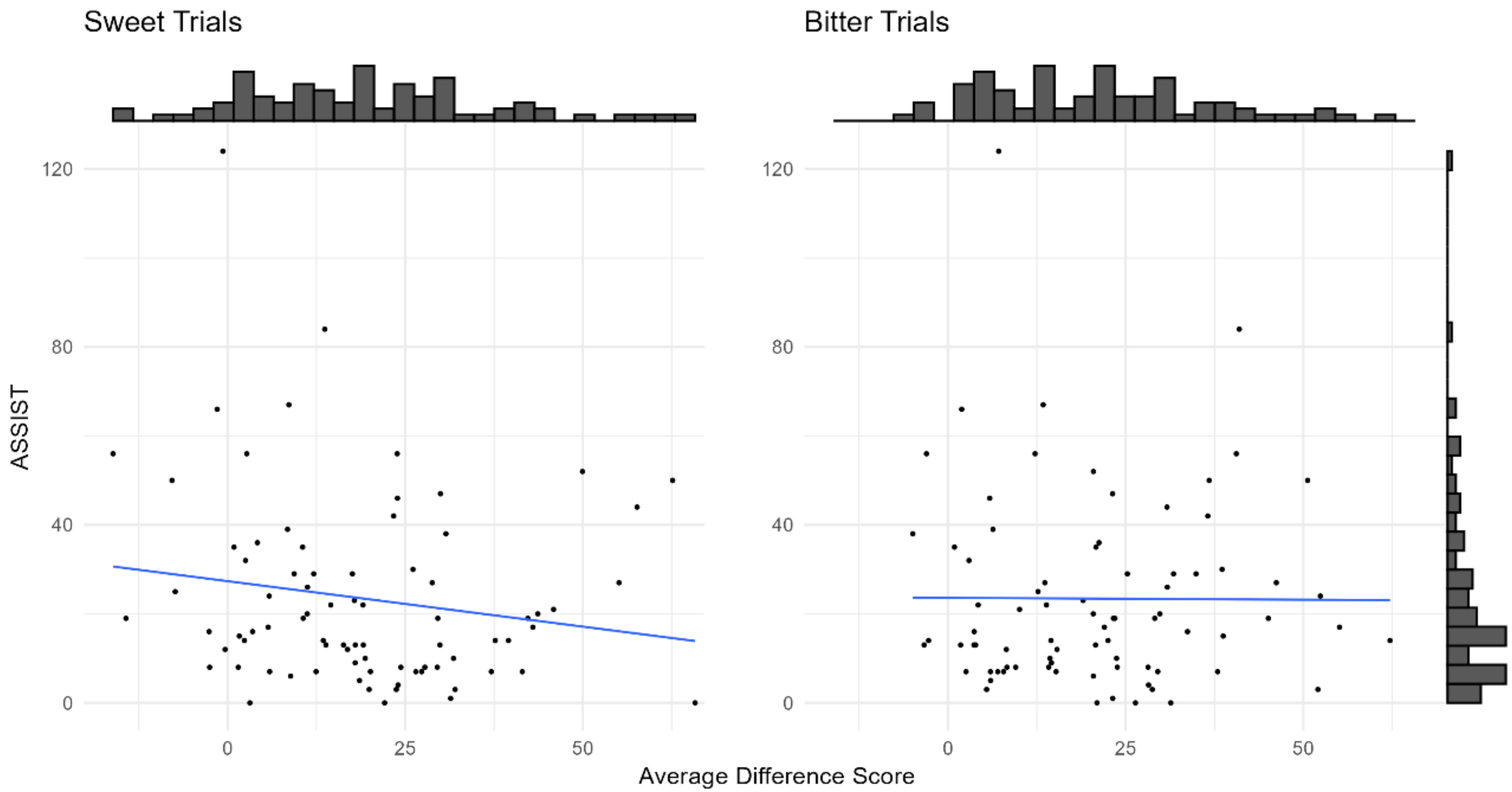
Note: * Significant at the $p < .05$ level.

**Regression model statistics**

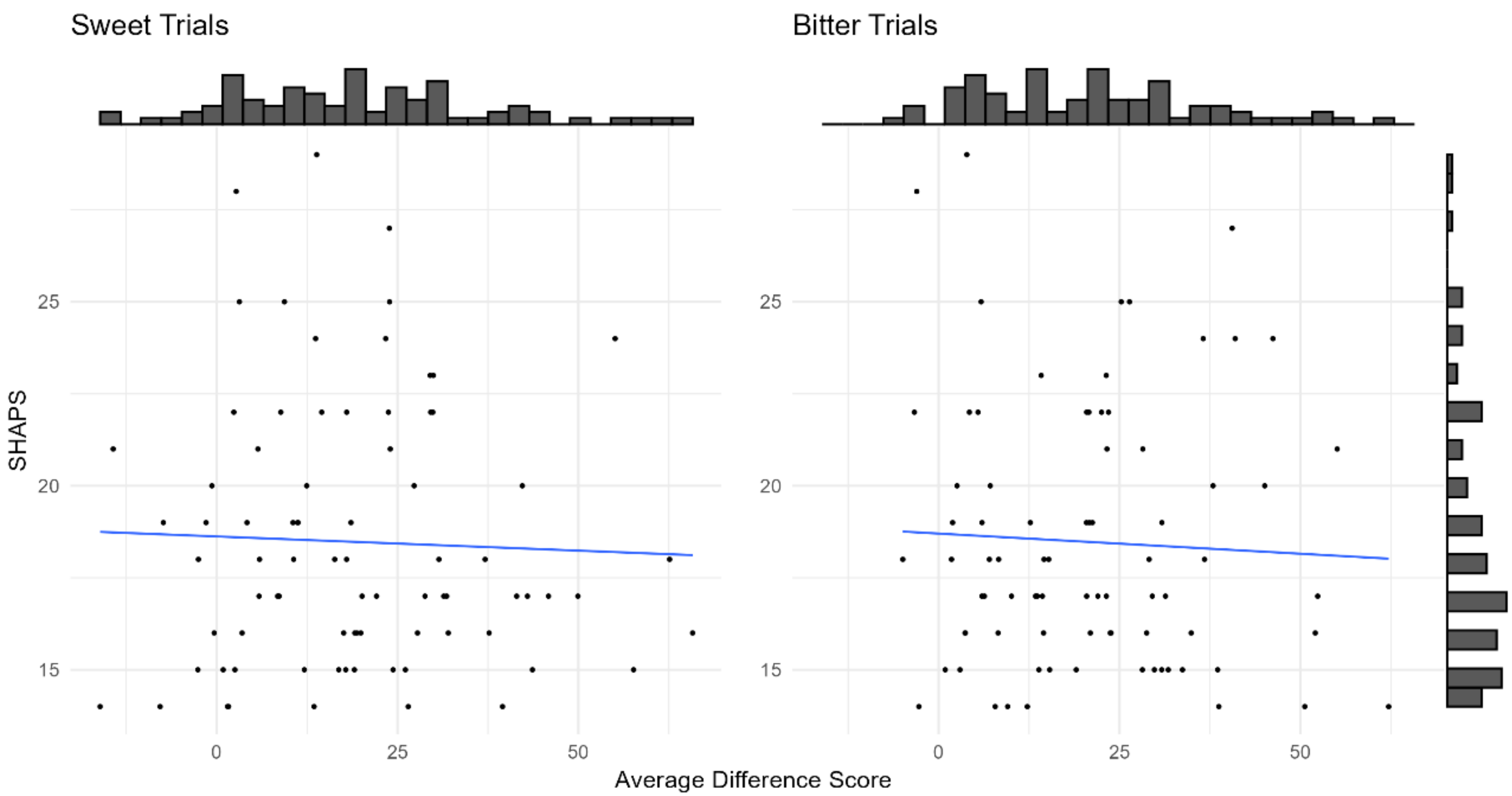
Standardized predicted values were plotted against standardized residuals to check for normality, linearity and homoscedasticity. The personality model showed no consistent deviation from zero, but in the psychopathology model, positive predicted values showed higher variance in residuals than negative predicted values, suggesting some degree of heteroscedasticity. However, a Breusch-Pagan test for heteroscedasticity did not indicate this was significant, $BP(2) = 1.56$, $p = .459$. Furthermore, a Shapiro-Wilk test of the residuals did not indicate significant deviation from normality, $W = .98$, $p = .248$. The maximum Cook's distance for the personality model was 0.21, which was much larger than the second largest Cook's distance of 0.07. Removing this case increased B for the UPPS-P, though not to significance, $p = .074$. The maximum Cook's distance for the psychopathology model was 0.08, which was in line with the next largest values, showing no disproportionate impact of individual cases.




*Supplementary Figure 2: Graph of UPPS-P scores against average differences between CS+ and CS- trials, separated by sweet and bitter trials. Histograms show marginal distributions. Trend lines show linear models.*




*Supplementary Figure 3: Graph of ASSIST scores against average differences between CS+ and CS- trials, separated by sweet and bitter trials. Histograms show marginal distributions. Trend lines show linear models.*




*Supplementary Figure 4: Graph of SHAPS scores against average differences between CS+ and CS- trials, separated by sweet and bitter trials. Histograms show marginal distributions. Trend lines show linear models.*

## BSI Subscales

The effect of the BSI was further investigated by analyzing its subscales. An average of the Variance Inflation Factors (VIF) was computed to assess multicollinearity, which showed that the subscales were moderately correlated with each other, average VIF = 3.22. This means a regression analysis may produce misleading results (Daoud, 2017). Therefore, Pearson correlation coefficients between the learning asymmetry and the BSI subscales are reported in Supplementary Table 6. Only the Phobic Anxiety subscale was significantly correlated with learning asymmetry, indicating that higher scores on this subscale were associated with a more negative learning asymmetry. However, the subscales with the next highest correlation coefficients, namely Depression and Anxiety, were only slightly lower compared to Phobic Anxiety. Given that there was a moderate degree of multicollinearity in addition to this, it would not be justified to interpret the contribution of Phobic Anxiety as meaningfully higher than other subscales.

**Supplementary Table 6**

*Bivariate correlations between the learning asymmetry and BSI subscales*

| | Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Learning asymmetry | — | | | | | | | | |
| 2 | Somatization | -.139 | — | | | | | | | |
| 3 | Obsession-Compulsion | -.124 | .622** | — | | | | | | |
| 4 | Interpersonal Sensitivity | -.184 | .560** | .549** | — | | | | | |
| 5 | Depression | -.214 | .509** | .710** | .680** | — | | | | |
| 6 | Anxiety | -.209 | .596** | .750** | .583** | .702** | — | | | |
| 7 | Hostility | -.138 | .525** | .546** | .542** | .686** | .577** | — | | |
| 8 | Phobic Anxiety | -.259* | .415** | .505** | .440** | .604** | .693** | .428** | — | |
| 9 | Paranoid Ideation | -.187 | .470** | .487** | .585** | .707** | .571** | .698** | .596** | — |
| 10 | Psychoticism | -.105 | .550** | .662** | .732** | .815** | .635** | .535** | .507** | .689** |

Note: * Significant at the $p < .05$ level. ** Significant at the $p < .01$ level.



.

**Learning asymmetry extremeness**

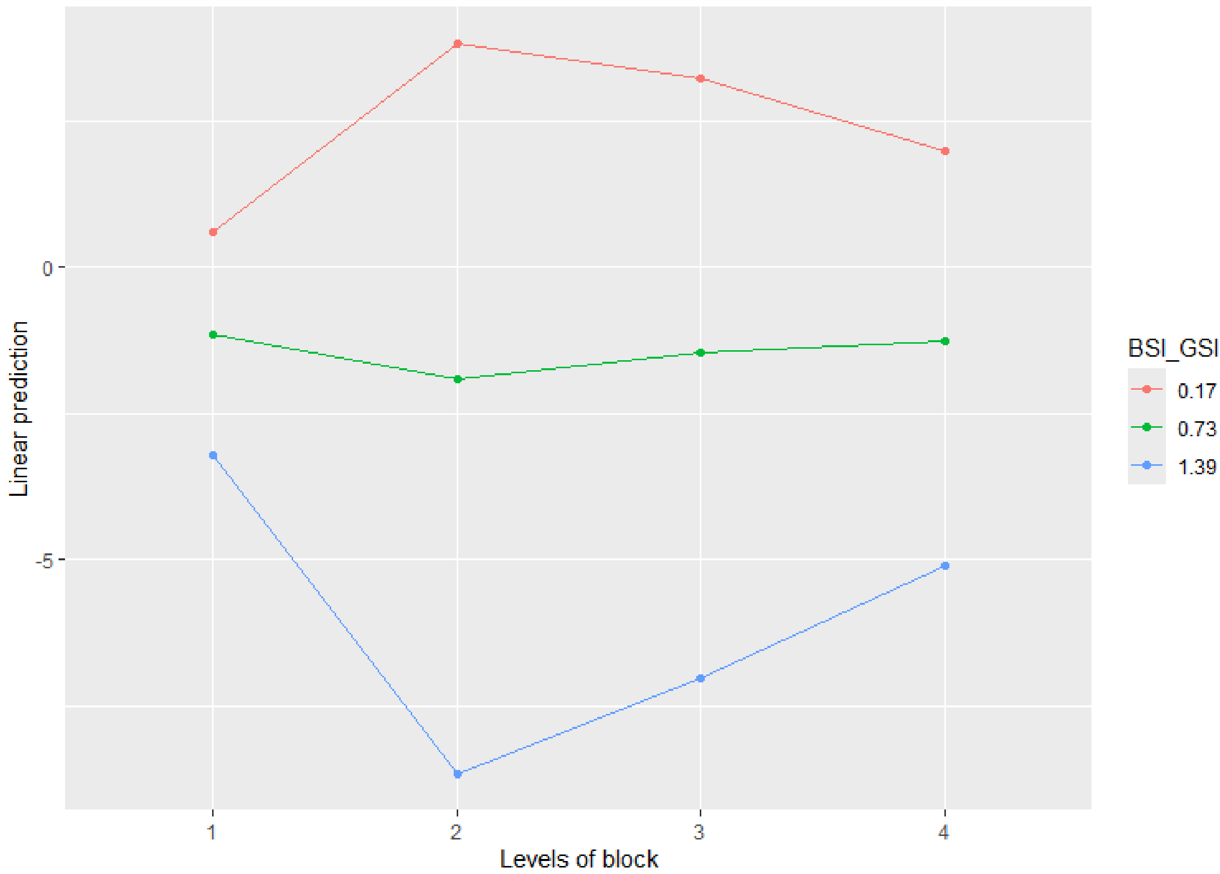
Although learning asymmetry is intended to measure the bias towards appetitive or aversive learning independent of the objective strength of learning, it is still possible that higher differences between CS+ and CS- resulted in corresponding higher learning asymmetry values. To test this, we determined how far from complete uncertainty each expectancy rating was by subtracting 50 from each and taking the absolute value. We then calculated the mean extremeness of each participant. This created a value that was higher the more the participant gave ratings closer to 0 or 100. Then we calculated the correlation between extremeness, learning asymmetry, and absolute learning asymmetry to determine whether confidence correlated with either the direction or the magnitude of learning asymmetry. This showed that the correlation between learning asymmetry and extremeness was not significant, $r(78) = -.078$, $p = .493$, but the correlation between absolute learning asymmetry and extremeness was significant, $r(78) = .322$, $p = .004$. Furthermore, the correlation between learning asymmetry and absolute learning asymmetry was significant, $r(78) = -.222$, $p = .048$. This shows that participants who rated their expectancy closer to 0 or 100 also had a learning asymmetry value further from 0, but were not significantly more likely to have a positive or negative learning asymmetry.

Overall learning strength was then determined by calculating the mean of the expectancy rating difference between CS+ and CS- for blocks 2, 3 and 4, also averaging sweet and bitter blocks. There was a significant correlation between overall learning strength and absolute learning asymmetry, $r(78) = .268$, $p = .016$, but not between overall learning strength and learning asymmetry, $r(78) = .113$, $p = .317$. This shows that participants who showed a larger difference between expectancy ratings for CS+ and CS- trials also showed a larger learning asymmetry. The correlation between overall learning strength and extremeness was also significant, $r(78) = .362$, $p < .001$.

To determine whether this influenced the significant correlation between learning asymmetry and psychological distress, we also performed correlations between the extremeness of the expectancy ratings split by taste (sweet vs. bitter). While the correlation between BSI and extremeness on sweet trials was marginally significant, $r(78) = .215$, $p = .056$, and so was the correlation between BSI and extremeness on bitter trials, $r(78) = .217$, $p = .053$, we also performed Steiger's Z test controlling for the correlation between extremeness on sweet and bitter trials, $r(78) = .881$, $p < .001$, and this test was not significant, $Z_H(77) = -.05$, p = .959. This shows that, while learning asymmetry in general may have been influenced by individual differences in expectancy rating extremeness, these were not differentially associated with appetitive versus aversive learning.

**Estimated Marginal Means of BSI: Learning Asymmetry per Block**

In the mixed-effects model using BSI and Block as fixed effects, with Block being analyzed as a factor, estimated marginal means were calculated to determine the size of the BSI effect at each point during the task. BSI scores were binned with the mean score across all participants as a medium level and one standard deviation above and below the mean as the low and high level (see Supplementary Table 3). Per-block contrasts showed that there were no significant effects of BSI in the first block, $t(249) = 0.69$, p = .771, third block, $t(249) = 1.86$, p = .153, or fourth block, $t(249) = 1.28$, p = .407, but there was a marginally significant effect in the second block, $t(249) = 2.26$, p = .063 (see Supplementary Figure 5). This suggests that the learning asymmetry effect was largest when participants had only one opportunity to learn the associations, and that the effect decreased as their learning improved, a similar result as was found in Kemp et al. (2024).




*Supplementary Figure 5: Estimated marginal means for learning asymmetry values per block and per BSI score bin. Red: low scores, green: medium scores, blue: high scores.*

# Chapter 4

The Balance of Appetitive and Aversive Learning in Personality and Psychopathology

Submitted as: Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. The balance of appetitive and aversive learning in personality and psychopathology. *Journal of Experimental Psychopathology.*



.

**Abstract**

Most conditioning studies on mental disorders focus solely on either appetitive or aversive learning, yet both may be essential to understanding the development of mental disorder symptoms. This study examines the relative strength of appetitive versus aversive learning – referred to as the learning asymmetry – in relation to anxiety, impulsivity, food reactivity, and symptoms of mental disorders. 101 volunteers participated in learning to associate complex objects with either gains or losses and made decisions to gain points or avoid losing them. We hypothesized that the strength of learning asymmetry would correlate with personality traits and symptoms related to appetitive or aversive learning. Stronger appetitive learning was associated with stronger food reactivity, weaker aversive learning was related to stronger impulsivity, and both stronger appetitive and weaker aversive learning were associated with higher trait anxiety. Other symptoms of mental disorders were however not significantly related to the learning asymmetry. While the associations suggest potential links between learning asymmetry and personality traits, the results were not conclusive enough to indicate that learning asymmetry alone can explain individual differences in mental disorder symptoms. Nonetheless, the current study provides valuable insights into individual differences in learning processes and how they relate to psychopathology.

## Introduction

In human learning, pleasant and unpleasant experiences influence future behavior, but not all individuals learn equally from each experience. Individual differences in learning can result in maladaptive behavior: oversensitivity to expected appetitive outcomes may cause overly risky behavior while oversensitivity to anticipated aversive outcomes may cause one to miss valuable opportunities. Such an imbalance between appetitive and aversive learning may play a role in the development of psychopathology including addictive disorders, mood disorders and anxiety disorders. As such, studying appetitive and aversive learning may help us understand how differences in learning influence mental health.

In conditioning studies, individuals learn the association between a neutral stimulus (i.e., conditioned stimulus; CS) and a rewarding, appetitive stimulus or an unpleasant, aversive stimulus (i.e., unconditioned stimulus; US). Conditioned stimuli are usually contrasted between those that are paired by a US (CS+) and those that are not (CS-). Conditioning studies have investigated how appetitive and aversive learning relate to maladaptive behavior in response to reward and punishment, and whether this is associated with the maintenance of mental disorders. A long-standing area of research explores how fear and anxiety disorders persist, often through mechanisms like heightened responsiveness to aversive stimuli (Pitman & Orr, 1986) . Several meta-analyses show that anxiety disorder patients have worse CS+/CS- discrimination driven by increased fear responses to the CS- compared to healthy controls, suggesting a tendency to engage in unwarranted fear responses (Duits et al., 2015; Lissek et al., 2005). By contrast, flying phobia patients show larger, not smaller, differences in acquisition between aversive and neutral stimuli compared to controls, if the stimuli used are more ambiguous (Vriends et al., 2012). Relatedly, men at risk for alcoholism are less able to discriminate between aversive and neutral stimuli (Finn et al., 1994), similar to what is seen in anxiety disorder patients.

Other studies on conditioning have primarily explored appetitive learning and its relation to psychopathology, particularly focusing on the distinction between an appetitive CS+ and a neutral CS-. However, the research in this area is relatively limited compared to the extensive work on aversive learning and anxiety. Within the field of eating disorders and obesity research, some studies demonstrate that overweight/obese individuals show poor discrimination of appetitive and neutral stimuli relative to those with healthy weight (Coppin et al., 2014; van den Akker et al., 2017; Zhang et al., 2014), yet others find the reverse, namely better CS+/CS- discrimination (Meemken et al., 2018; Meyer et al., 2015) or report no learning differences depending on weight (van den Akker et al., 2019).

Advancing our understanding of how individual differences in appetitive and aversive conditioning relate to psychopathology is hindered by the scarcity of studies that examine both types of conditioning, rather than focusing on just one in isolation. In a study using electric shocks and monetary rewards, Zbozinek et al. (2021) showed that trait anxiety was associated with stronger aversive learning and that higher trait depression was associated with weaker appetitive learning, suggesting that a concurrent approach can improve our understanding of how conditioning and psychopathology interact. Unfortunately, very few studies have systematically applied such an approach, and embedding this research in a theoretical framework is crucial to developing more insight into the underlying mechanisms.

The question of how appetitive and aversive learning may result in maladaptive behavior has been investigated through Gray's Reinforcement Sensitivity Theory (RST, Gray, 1976). This theory originally sought to explain processes of emotion, learning and motivation through neural systems governing adaptive behavior in response to competing goals, but over time its proposed mechanisms have become centered around personality traits that reflect individual sensitivity to appetitive and aversive learning (Smillie et al., 2006). These traits have been described as impulsivity, inhibition[3], and fear, which are governed by the Behavioral Activation System (BAS), the Behavioral Inhibition System (BIS) and the Fight, Flight, Freeze System (FFFS) respectively: the FFFS responds to signals of immediate danger (i.e. by fear), the BAS responds to (potential) reward signals (i.e. by impulsivity), and the BIS resolves goal conflict between the two by adjusting appetitive and aversive impulses or behaviors (i.e. by inhibition; Pickering & Corr, 2008). Individuals can vary in how sensitive they are to appetitive and aversive learning, which should be reflected in the impulsivity and inhibition traits. Importantly, disorders such as anxiety or panic disorder may be explained by the activation of these systems at

[3] Also referred to as anxiety, but inhibition is used here to distinguish it from other measures of anxiety.

“maladaptive intensity" (Pickering & Corr, 2008), and such a response may be caused by oversensitivity to appetitive or aversive stimuli. Investigating relative differences in appetitive and aversive learning and whether they correspond with predictions from RST may therefore clarify their role in the persistence of mental disorders.

While conditioning studies often include ‘neutral’ trials where no unconditioned stimulus is presented, RST views the absence of an expected appetitive or aversive stimulus as functionally equivalent to the presence of an aversive or appetitive stimulus, respectively (Gray, 1975). When a population of interest shows greater differences in learning between an aversive CS+ and a neutral CS- compared to a control group, this could imply that this group has either a higher sensitivity to aversive learning (i.e., learning about the CS+) or a lower sensitivity to appetitive learning (i.e., learning about the CS-). Given that differences in how well individuals discriminate between CS+ and CS- is the source of inconsistencies in several studies on anxiety disorders as well as overweight and obesity, it may be that how individuals perceive the contrast between appetitive/aversive and neutral stimuli affects how they respond to the CS+ and CS- in experimental settings. By measuring the relative differences in learning, we may clarify the role of appetitive and aversive learning in the persistence of mental disorders.

In earlier studies, we explored individual differences in appetitive and aversive learning using a conditioning task with primary and secondary reinforcers. Kemp et al. (2024) used a task where participants responded to stimuli based on the potential to gain or lose points, with these points linked to an additional monetary reward. Based on these responses, the difference in performance between gain (i.e., appetitive learning) and loss trials (i.e., aversive learning) was calculated, representing the learning asymmetry, and then linked to personality traits often associated with certain mental disorders. Results from Kemp et al., (2024) showed that learning asymmetry was significantly associated with a self-report measure of impulsivity, but not with neuroticism or anhedonia. The results indicated that high impulsive individuals performed worse than low impulsive individuals in aversive learning, without showing any differences in appetitive learning. Additionally, Kemp et al. (under review) used the same task with sweet and bitter tastes as the unconditioned stimuli, and in this study, the learning asymmetry was linked to psychological distress.



.

Specifically, participants with weaker appetitive learning showed higher psychological distress, but aversive learning showed no association with distress. As in Kemp et al. (2024), anhedonia showed no relation to learning asymmetry, but neither did impulsivity, which contrasts with the previous results. Whether learning asymmetry relates specifically to certain risk factors or symptoms of psychopathology, or merely psychological distress in general, is yet unclear.

The current study has two aims: one, to extend the results from the previous two studies into the potential for learning asymmetry to serve as a transdiagnostic factor for mental disorders, and two, to investigate whether learning asymmetry is compatible with RST, which may clarify the inconsistent results from previous conditioning studies. This study uses the same method as Kemp et al. (2024), namely a conditioning task using abstract 3D objects as conditioned stimuli and a system of point gain and loss, tied to an additional monetary reward, as unconditioned stimuli. The difference between the strength of gain and loss associations is compared to various self-report measures related to personality traits and mental disorders. We hypothesized that personality traits and mental disorder symptoms would relate to learning asymmetry as follows: a more positive learning asymmetry would be associated with higher impulsivity and higher food reactivity, while a more negative learning asymmetry would be associated with more symptoms of depression and anxiety, and higher inhibition. Based on previous research, we specifically expect weaker aversive learning to be associated with higher impulsivity, and weaker appetitive learning to be associated with higher psychological distress.

## Method

### Participants

101 undergraduate students participated in the lab experiment. They were recruited through the university's research participation system for psychology students. Participants were between 18 and 26 years of age, $M = 20.22$, $SD = 1.69$, and were 89.1% female. Participants' national origin was predominantly German (43.6%) and Dutch (24.8%) with other European nationalities accounting for 22.8% of the sample and non-European nationalities for 8.9%. Testing occurred between 26

January 2024 and 26 March 2024. Participants took part in the study for either study credit or financial compensation. Furthermore, participants who scored above average on the learning task were given a €5 voucher in addition to their chosen compensation.

## Materials

### *Stimuli*

Stimuli were generated using the scripts and instructions provided by Watson et al. (2019) on 3DS Max 2022. Twenty-six stimuli were generated, ensuring that no two stimuli shared more than one feature between them. Thirteen stimuli were associated with point gain and thirteen with point loss, with a 100% reinforcement rate for each. These associations were counterbalanced between participants, and assigned such that no single feature was consistently associated with gain or loss. (See Figure 1)

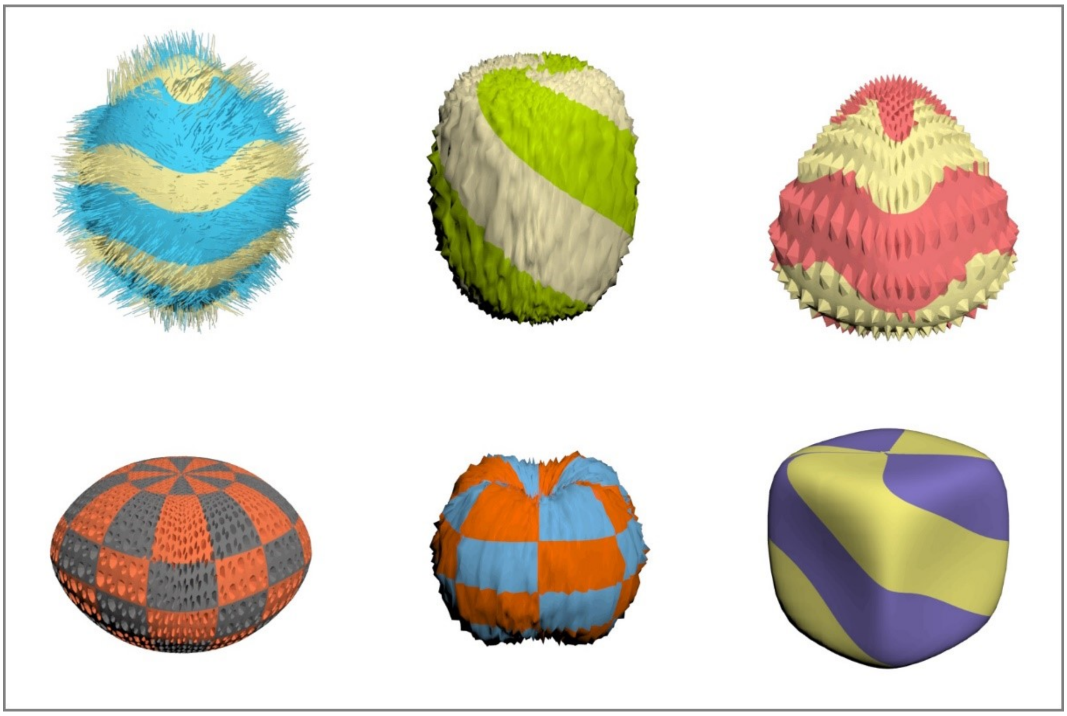


*Figure 1: Six examples of Quaddles used as conditioned stimuli.*

### *Task Performance*

After completing the task, participants received nine questions about the task, to verify that they completed the task as intended and that no problems occurred with the online test platform. Details on these questions are available in Supplementary Table 1.

### *Brief Symptom Inventory (BSI)*

The BSI is a self-report scale for psychological distress. It consists of 53 items rated on a 5-point scale, assessing how much participants were distressed by a given symptom during the past seven days, ranging from 0 (Not at all) to 4 (Extremely). The items cover nine symptom dimensions: Somatization, Obsession-Compulsion, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. The Global Severity Index (GSI) is calculated as the mean of all 53 items. Cronbach's alpha for this scale was very high ($\alpha$ = .960).

### *UPPS-P Impulsive Behavior Scale*

The UPPS-P Impulsive Behavior Scale is an inventory of various kinds of impulsive behaviors (Lynam et al., 2006), consisting of 59 items scored on a scale of 1 (Agree Strongly) to 4 (Disagree Strongly). The subscales include (Negative) Urgency, (lack of) Premeditation, (lack of) Perseverance, Sensation Seeking, and Positive Urgency. To obtain the total impulsivity score, all subscale scores are averaged. Cronbach's alpha for this scale was very high ($\alpha$ = .926).

### *The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)*

The ASSIST is a substance abuse screening test that measures the frequency of recreational use of various substances, as well as the degree to which this use results in problems (WHO, 2002). It consists of a list of ten categories of substances, and participants are asked if they ever used this substance recreationally. For each category that they answered in the affirmative, they are then asked six questions about their use of this substance and its impact on their life (e.g. "How often have you used cannabis?"), which are answered in the form of a five-point scale ranging from "Never" to "Daily or almost daily", or on a three point scale with the choices "No, never", "Yes, in the past 3 months", and "Yes, but not in the past 3 months". More frequent use and more risky substances are reflected in higher scores. The sum of all items was used as the total score. Cronbach's alpha for this scale was high ($\alpha$ = .878).

### *The Behavioral Inhibition Scale/Behavioral Activation (BIS/BAS) scale*

The BIS/BAS scale is a 20-item self-report scale measuring approach and avoidance behavior tendencies. It contains three BAS subscales (Drive, Fun Seeking, and Reward Responsiveness) and one BIS subscale. The items are answered on a four-point scale ranging from 4 (Disagree strongly) to 1 (Agree strongly). Cronbach's alpha for these scales ranged from acceptable ($\alpha_{BIS}$ = .740; $\alpha_{BAS-D}$ = .771) to middling ($\alpha_{BAS-F}$ = .559; $\alpha_{BAS-R}$ = .547).

### *The Depression, Anxiety, Stress Scale (DASS)*

The DASS is a 21-item self-report scale designed to measure the severity of a range of symptoms common to both depression and anxiety. For each item, the participants answers how much this applied to them over the past week, ranging from 0 (Not at all) to 3 (Almost always). Cronbach's alpha for this scale was very high ($\alpha$ = .933).

### *The Power of Food Scale (PFS)*

The PFS is a 15-item self-report scale measuring the hedonic value of food and its impact on behavior. Items were answered on a five-point scale ranging from 1 (Don't agree at all) to 5 (Strongly agree). Subscales include Food Available, Food Present, and Food Tasted. To obtain the total score, all subscales were averaged. Cronbach's alpha for this scale was very high ($\alpha$ = .906).

## Procedure

The experimental procedure was approved by the local Ethics Review Committee Psychology and Neuroscience (approval code OZL_234_27_02_2021_S7_A1). Participants were invited to the lab and seated at a computer in a soundproofed room, where they were provided with an information letter and signed the informed consent form. They were then verbally instructed on the details of the task, including that they would receive an additional reward if they scored more than a certain number of points. This additional reward was given so that performing well on the task would result in a real reward rather than being limited to a point score, which emphasized the appetitive and aversive associations created by the task. The first block (the acquisition block) showed participants the association of each of the twenty stimuli; no responses were necessary. The subsequent three blocks (the response blocks) repeated the same stimuli and asked participants to respond to each stimulus based on the associations they learned from the acquisition block.

The acquisition block indicated the point value of each stimulus ("Gains 10 points"/"Loses 10 points"). During the subsequent three response blocks the participants were asked to 'accept' or 'reject' the outcome of each stimulus. The goal was to gain points and avoid losing points by 'accepting' stimuli associated with point gain and 'rejecting' stimuli associated with point loss. This was done by clicking one of two buttons, 'Yes' or 'No', when given the question "Accept the result?". On trials where the stimulus signaled point gain (hereafter: 'Gain trials'), responding 'Yes' would add 10 points. Conversely, on trials where the stimulus signaled point loss (hereafter: 'Loss trials'), responding 'Yes' would subtract 10 points. Responding 'No' resulted in neither point gain nor loss. In this way, Gain trials created appetitive associations through the possibility of gaining points, and Loss trials created aversive associations through the possibility of losing points.

After responding to each trial, participants selected a point on a 100-point visual analog scale to indicate their confidence in their decision. They then received feedback indicating the value of the stimulus. In addition, participants' point total for the entire task was displayed with each instance of feedback. For more details, see Figure 2.

Upon task completion, participants filled out the UPPS-P, the BSI, the ASSIST, the BIS/BAS scale, the DASS, and the PFS using the Qualtrics online survey platform. Participants were thanked and fully debriefed at the end of the study and received money or course credit as remuneration for their participation.

### Outcome measures

From the measures of accuracy and confidence, a composite variable was created using the sum of confidence ratings on all correct trials minus the sum of confidence ratings on all incorrect trials, then divided by the total number of trials, a value that is henceforth referred to as 'weighted accuracy' or 'wAcc'. By weighting participants' answers according to their confidence, wAcc more closely represents the strength of the learned associations.

### Transparency and Openness

The study's hypotheses, analyses, exclusion criteria and sample size were preregistered at https://aspredicted.org/WHY_KR5. All data, analysis code, stimulus materials, and task script are available at https://osf.io/27f4w/?view_only=536019eb37f143f9a82824fcd5edbf2f.

At preregistration, it was determined that outliers on the learning asymmetry and on self-report measures exceeding 2.5 standard deviations from the mean would be winsorized.

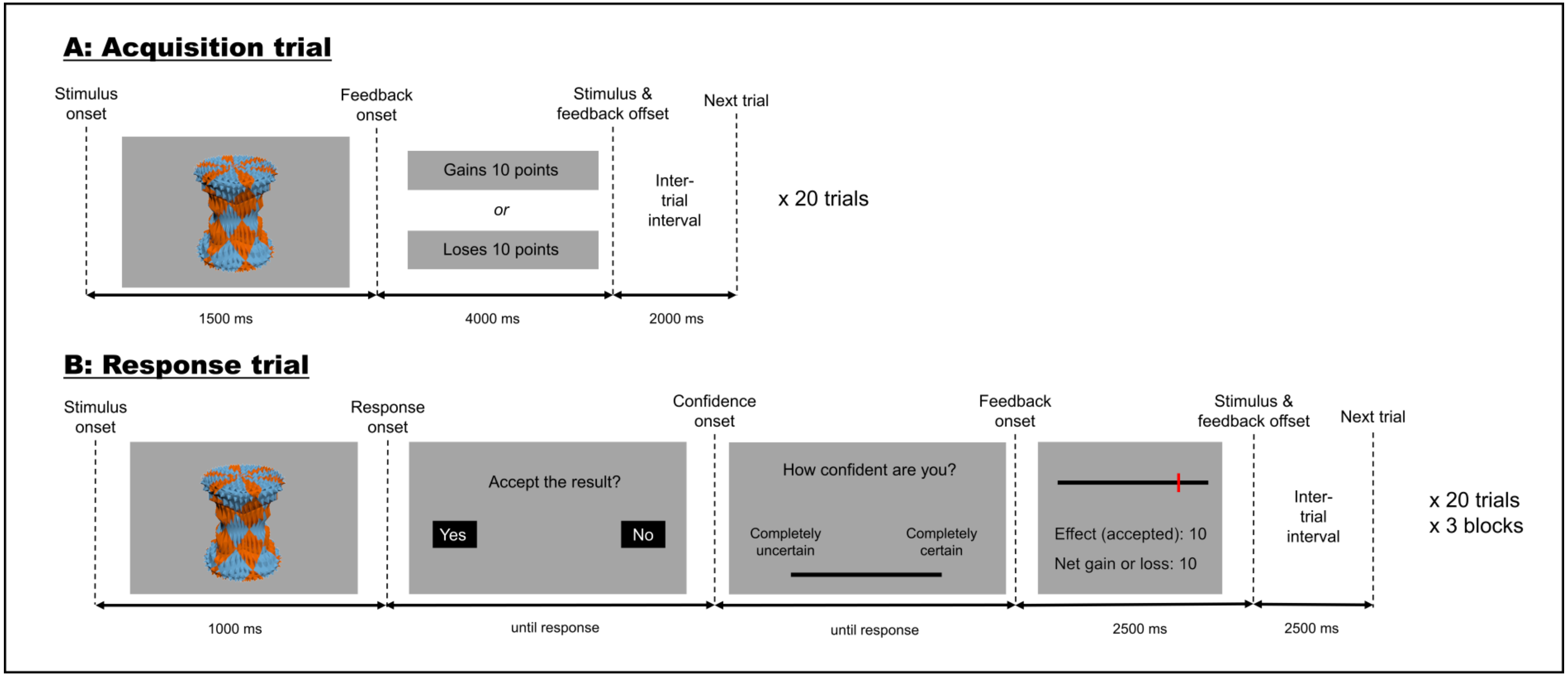



*Figure 2: Diagram of the task procedure for the learning task, split by Acquisition trials (A) and Response trials (B). Areas in grey represent elements displayed during the task. The feedback shown for the response trial is an example of a Yes response to a Gain stimulus.*

# Results

## Inclusion and Exclusion

First, self-report measures were examined, and found to contain several values that deviated more than 2.5 standard deviations from the mean. At preregistration, these values were planned to be winsorized, as in a previous study, certain outliers on self-report measures were found to disproportionally affect the results. However, the threshold of 2.5 standard deviations was found to be overly conservative for the current study and would remove too much variation from the sample. Therefore, the self-report measures were not winsorized. Second, although there were no wAcc asymmetry values meeting the preregistered criteria for outliers, several extreme values were present nonetheless. Such extreme values may have been the result of participants responding with ‘Yes’ or ‘No’ disproportionally compared to the ideal ratio of 1:1, indicating poor adherence to the task requirements. In accordance with Kemp et al. (2024), we determined the mean proportion of ‘Yes’ responses relative to total responses for the entire sample ($M = .485$, $SD = .099$) and excluded participants who deviated more than two standard deviations from the mean. Six participants met these criteria and were excluded. An analysis including these participants can be found on pp. 106.

## Task Performance

To determine whether learning improved over time, a repeated-measures analysis of variance (ANOVA) was performed using Valence (Gain vs. Loss) and Block (1 vs. 2 vs. 3) as factors and wAcc as the dependent variable. No effect of Valence was found, $F(94) = 1.07$, $p = .304$, but the effect of Block, $F(94) = 70.82$, $p < .001$, and the interaction between Block and Valence, $F(94) = 9.87$, $p < .001$, were significant. To unpack these effects, paired t-tests were performed, using the Holm-Bonferroni correction to determine the significance threshold. These showed that the difference between wAcc on Gain and Loss trials was significant for blocks 1, $t(94) = -2.79$, $p = .006$, and 3, $t(94) = 2.53$, $p = .013$, but not for block 2, $t(94) = 1.67$, $p = .098$. Gain trials showed higher wAcc than Loss trials on block 1, but the reverse in block 3 (see Figure 3). Furthermore, differences in wAcc were not significant for gain trials between block 1 and 2, $t(94) = -2.05$, $p = .043$, but they were significant for gain trials between block 2 and 3, *t*(94) = 4.53, *p* < .001, for loss trials between block 1 and 2, *t*(94) = 3.91, *p* < .001, and for loss trials between block 2 and 3, *t*(94) = 7.34, *p* < .001, with wAcc increasing in every block with the exception of Gain trials between block 1 and 2 (see Figure 3). Therefore, although wAcc on Gain trials increased more slowly, wAcc for both Gain and Loss trials significantly increased by the final block of the task, indicating that learning improved as the pairings were repeatedly presented.

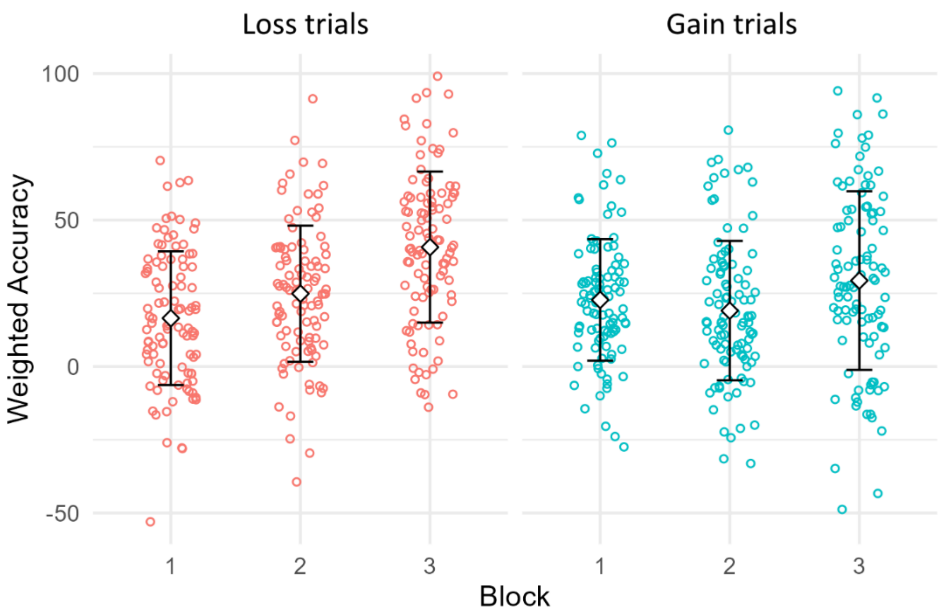



*Figure 3: Comparison of weighted accuracy by valence (Gain vs. Loss) and block. Circles indicate participant means. Diamonds indicate sample means. Error bars indicate sample standard deviations.*

**Learning Asymmetry and Self-Report Measures**

Descriptive statistics for the self-report measures can be found in Supplementary Table 1. The wAcc asymmetry did not significantly differ from zero according to a one sample t-test, *t*(94) = -1.03, *p* = .304. A Kolmogorov-Smirnov test showed that the distribution did not significantly deviate from normality, *D*(95) = 0.066, *p* > .200. Two linear regression analyses were performed using the wAcc asymmetry as the dependent variable. First, the UPPS-P, the BIS, the three BAS subscales, and the PFS were used as predictors, as they measure personality traits relevant to appetitive and aversive learning. This model including all predictors explained a significant amount of variance, $F(6, 88) = 2.78$, $p = .016$, $R^2 = .159$. All three BAS subscales did not meet the inclusion criterion and were excluded. The final model including the UPPS-P, the PFS, and the BIS, was significant, $F(3, 91) = 4.46$, $p = .005$, $R^2 = .131$. While the BIS met the inclusion criterion, it was only marginally significant as a predictor ($p = .058$) while the other predictors were significant ($p$s < .038).

For the second regression analysis, the General Severity Index (GSI) of the BSI, the sum score of the ASSIST, and the three subscales of the DASS were used as predictors, as they measure symptoms relevant to mental disorders. This model did not explain a significant amount of the variance, $F(5, 89) = 1.17$, $p = .333$, $R^2 = .061$. The BSI, the ASSIST and the Stress subscale of the DASS did not exceed the removal criterion and were excluded. The final model consisted of the Anxiety and Depression subscales of the DASS, $F(2, 92) = 2.47$, $p = .091$, $R^2 = .051$. Anxiety met the inclusion criterion but was marginally significant ($p = .059$) and Depression was significant as a predictor ($p = .037$). Additional model statistics can be found in Table 1.

These results show that several predictor variables were associated with learning asymmetry, but they do not indicate whether appetitive or aversive learning plays the main role in these effects. Therefore, Pearson correlation coefficients between the wAcc on Gain and Loss trials and the predictors were compared to determine the association with Gain and Loss wAcc. The PFS was mainly associated with higher Gain trial wAcc, the UPPS-P was mainly associated with lower Loss trial wAcc, DASS-Depression was mainly associated with higher Loss trial wAcc, and BIS was associated with both higher Gain trial wAcc and lower Loss trial wAcc. Correlation coefficients between the included predictor variables and corresponding significance level are reported in Table 2. Correlation coefficients between all predictors can be found in Supplementary Table 3.

**Table 1**

*Summary of backwards entry regression analysis for personality traits and mental disorder symptoms*

| Model | Predictors | *B* | SE *B* | *β* | SE *β* | *t* | *p* | 95% CI for B | | Tolerance |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper | |
| 1 | **UPPS-P** | **17.39** | **7.88** | **.28** | **.13** | **2.21** | **.030** | 1.7 | 33.1 | .60 |
| | **PFS** | **7.11** | **2.88** | **.26** | **.11** | **2.47** | **.016** | 1.4 | 12.8 | .85 |
| | **BIS** | **11.60** | **5.17** | **.24** | **.11** | **2.24** | **.028** | 1.3 | 21.9 | .83 |
| | BAS-D | -0.58 | 4.25 | -.01 | .11 | -0.14 | .892 | -9.0 | 7.9 | .86 |
| | BAS-F | 6.38 | 5.64 | .15 | .13 | 1.13 | .261 | -4.8 | 17.6 | .56 |
| | BAS-R | -9.87 | 6.43 | -.17 | .11 | -1.54 | .128 | -22.6 | 2.9 | .81 |
| 1′ | **UPPS-P** | **13.17** | **6.27** | **.21** | **.10** | **2.10** | **.038** | 0.7 | 25.6 | .95 |
| | **PFS** | **6.34** | **2.79** | **.23** | **.10** | **2.27** | **.025** | 0.8 | 11.9 | .91 |
| | BIS | 9.49 | 4.94 | .20 | .10 | 1.92 | .058 | -0.3 | 19.3 | .91 |
| 2 | BSI-GSI | -2.13 | 7.61 | -.06 | .20 | -0.28 | .780 | -17.3 | 13.0 | .26 |
| | ASSIST | -0.10 | 0.13 | -.10 | .12 | -0.78 | .436 | -0.4 | 0.2 | .69 |
| | DASS-Depression | -1.46 | 1.10 | -.27 | .20 | -1.33 | .186 | -3.6 | 0.7 | .25 |
| | DASS-Anxiety | 1.39 | 0.99 | .22 | .16 | 1.41 | .161 | -0.6 | 3.4 | .42 |
| | DASS-Stress | 0.79 | 1.03 | .14 | .18 | 0.77 | .443 | -1.3 | 2.8 | .34 |
| 2′ | **DASS-Depression** | **-1.56** | **0.74** | **-.29** | **.14** | **-2.12** | **.037** | -3.0 | -0.1 | .55 |
| | DASS-Anxiety | 1.63 | 0.85 | .26 | .14 | 1.91 | .059 | -0.1 | 3.3 | .55 |

Note: Model 1 = personality traits, model 2 = mental disorder symptoms, prime (′) = final model. Coefficients significant at the $p < .05$ level are bolded.

**Table 2**

*Bivariate Correlations for wAcc and Predictors Meeting the Inclusion Criterion*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. wAcc asymmetry | — | | | | | | |
| 2. Gain trial wAcc | .619** | — | | | | | |
| 3. Loss trial wAcc | -.438** | .431** | — | | | | |
| 4. DASS – Depression | -.117 | .006 | .139 | — | | | |
| 5. DASS – Anxiety | .076 | .083 | .019 | .671** | — | | |
| 6. BIS | .207* | .130 | -.108 | .399** | .516** | — | |
| 7. UPPS-P | .145 | .037 | -.114 | -.563** | -.322** | -.178 | — |
| 8. PFS | .249* | .259* | .010 | .224* | .314** | .272** | -.175 |

Note: *: $p < .05$. **: $p < .01$.

To determine whether wAcc for Gain or Loss trials played a larger role in the significant associations with learning asymmetry, Steiger's Z-test was performed on the correlation coefficients (Hoerger, 2013). This was not significant for Depression, $Z_H = 1.24$, $p = .215$, Anxiety, $Z_H = 0.6$, $p = .552$, or the UPPS-P, $Z_H = -1.41$, $p = .16$, but it was significant for the BIS, $Z_H = 2.22$, $p = .026$, and the PFS, $Z_H = 2.35$, $p = .018$.

In summary, the first regression analysis showed that impulsivity, food reactivity and inhibition were associated with wAcc asymmetry. There were several reasons for this: higher impulsivity was more strongly associated with aversive learning (lower wAcc for Loss trials); higher food sensitivity was more strongly associated with appetitive learning (higher wAcc for Gain trials), and higher inhibition was associated with both appetitive and aversive learning (higher wAcc on Gain trials and lower wAcc on Loss trials). However, in the final model, inhibition was only marginally significant, and the difference between correlation coefficients for Gain and Loss trials was not significant for impulsivity. The second regression analysis showed that depression and anxiety were associated with wAcc asymmetry. Higher depression was more strongly associated with lower wAcc on Loss trials than Gain trials and higher anxiety was more strongly associated with higher wAcc on



.

Gain trials than Loss trials, but neither correlation coefficient differed significantly between Gain and loss trials.

## Discussion

This study employed a conditioning task to measure the relative strength of appetitive and aversive learning using intermixed appetitive and aversive trials. Accuracy and confidence ratings were combined into a composite measure of learning, known as learning asymmetry, which was compared to self-report measures of personality and symptoms of psychopathology. The personality traits investigated included the BIS/BAS scale, a measure of reinforcement sensitivity; the UPPS-P, a measure of impulsivity; and the Power of Food Scale, a measure of the hedonic value of food. Conversely, psychopathological factors were investigated using the BSI, a measure of psychological distress; the DASS, a measure of depression, anxiety and stress; and the ASSIST, a measure of substance use.

The results showed that among the personality traits, the BAS subscales did not significantly predict the learning asymmetry, which is contrary to our hypotheses, but the UPPS-P, BIS and PFS did, indicating that impulsivity, inhibition and food sensitivity all play a role in the direction and strength of learning asymmetry. Interestingly, and opposite to what was hypothesized, inhibition was associated with both stronger appetitive learning and weaker aversive learning. This was, in contrast, not found for anxiety, which may indicate that the reinforcement in the current task was not experienced as particularly aversive, despite previous research having indicated that conditioning using monetary rewards yielded similar results as primary reinforcers (M. R. Delgado et al., 2006). Furthermore, both the UPPS-P and the BAS subscales are purported to measure impulsivity, but only the UPPS-P showed a relation to learning asymmetry. This may be explained by the two measuring different constructs: the use of 'impulsivity' for the trait measured by the BAS has been noted to be somewhat inappropriate given that it encompasses adaptive processes (Pickering & Corr, 2008), in contrast to the UPPS-P which measures maladaptive impulsive behavior. This is more consistent with the conception of learning asymmetry as a measure of disproportionate sensitivity to appetitive or aversive stimuli, but given that the same effect of the UPPS-P was not found in Kemp et al. (2025), it may not be the whole explanation. In said study, sweet and bitter tastes were used as unconditioned stimuli rather than point gain and loss, which necessitated the appetitive and aversive trials to be separated rather than intermixed. This difference in methodology may have affected the results, and emphasizes that the effects of study design on measures of appetitive and aversive learning warrants further research.

Finally, the Power of Food Scale, a measure of the hedonic value of food and its impact on the individual, was included to investigate the relevance of learning asymmetry to eating behavior. Given that our sample of volunteers did not specifically include participants with overweight, we investigated hedonic value as it has been linked to food intake beyond satiety (Levitsky & Shen, 2008). The PFS has not been measured in previous learning asymmetry studies, but showed a significant effect, indicating that participants with higher food reactivity show stronger appetitive learning, but no difference in aversive learning. This indicates that participants who are very reactive to food also learn better from Gain trials presented in the task. This supports the notion that this measure of food responsivity is in fact measuring a broader tendency towards appetitive learning including non-food stimuli, as suggested by Lowe et al. (2009). However, with no effect on aversive learning, this does not elucidate why appetitive conditioning is sometimes weaker (e.g. van den Akker et al., 2017) and sometimes stronger (e.g. Meemken et al., 2018) in overweight and obese individuals. Discovering the mechanisms of appetitive conditioning and food sensitivity will require further study.

The conclusions we may draw about learning asymmetry and psychopathology are limited by testing a convenience sample rather than patients diagnosed with mental disorders, as well as by the sample being mostly female, which may have over- or under-represented certain symptoms compared to a gender-balanced sample. A further aim of this study was to investigate learning asymmetry as a transdiagnostic factor, and while the results showed that depression was a significant predictor of learning asymmetry, the association with aversive learning was not significantly different from the association with appetitive learning. A previous study found that higher psychological distress was associated with weaker appetitive learning (Kemp et al., 2025), but this was not replicated in the current study. Unlike the previous study, tolerance values for the psychopathology predictors were very low, which affected the reliability of the results. Another possible factor is that learning asymmetry itself is more state-like than trait-like. Contrasting appetitive with aversive learning by taking the difference between the two measures may not be the optimal method of investigating this. This method increases the effect size if appetitive and aversive learning are both influenced by a certain trait, but in the previous and current studies, this has only been the case for the BIS. Although the current research into the relationship between conditioning and psychopathology disproportionately focuses on either appetitive or aversive learning, there may be a more suitable method to address this. However, maintaining the 'weak' situation (Lissek et al., 2006), in which stimulus associations are ambiguous, remains an important tool in measuring individual differences in appetitive and aversive learning.

The current study shows that learning asymmetry is moderately influenced with certain personality traits, and slightly influenced by symptoms of psychopathology. In addition, the pattern of appetitive and aversive learning differences does not fit the predictions made by RST. We conclude that individual differences in appetitive and aversive learning are relevant to personality and psychopathology in ways that are not well understood due to the rarity of studies measuring appetitive and aversive learning together, and further research using the combined approach is warranted. However, the specific method of using learning asymmetry as a composite measure has not shown that it contributes unique insight into learning behavior and psychopathology. By itself, it was not a reliable indicator of psychopathology and did not match up with the predictions of RST. Alternative approaches may be more productive in the study of individual differences in learning.

## Supplementary Material

### Task Performance

**Supplementary Table 1**

*Questions and Answers from the Task Follow-Up Questionnaire*

| Question | Answers |
|---|---|
| 1. Were the tasks and instructions clear and understandable? | a) Yes, I understood everything perfectly<br>b) Yes, I understood most of it<br>c) No, there were a lot of things I didn't understand<br>d) No, I didn't understand any of it |
| 2. Please elaborate on what you found unclear. | |
| 3. Did you put a good effort into giving accurate answers? | a) Yes, I did the best that I could<br>b) Kind of, I didn't try as hard as I could have<br>c) Not really, I wasn't paying much attention<br>d) No, I was just clicking randomly |
| 4. Did you use any memorization techniques or other methods that helped you respond accurately? If yes, please describe them. If no, you can leave this blank. | |
| 5. Which of these words match your feelings about the task? Please select all that apply. | ▪ Interesting ▪ Boring<br>▪ Frustrating ▪ Difficult<br>▪ Challenging ▪ Rewarding |



.

| | |
|---|---|
| 6. How distinguishable did you find the different objects? | a) They were all clearly different<br>b) They were somewhat similar but still distinguishable<br>c) They had a lot of similarities and I mixed some of them up<br>d) They were so similar I mixed them up constantly |
| 7. Are there any other thoughts or opinions you had about the tasks? Please list any that you can think of. | |

Note: Empty cells in the 'Answers' column indicate that participants were allowed to type in their own answer.

In the task-related follow-up questions, all participants indicated that the instructions were fully understandable and that they made at least a moderate effort to perform well (see Supplementary Table 2). When asked about the distinguishability of the stimuli, 53.69% of participants reported them to be mostly or very distinguishable, which is expected given that the stimuli consisted of many combinations of a few different properties. As such, these answers indicated that the balance between recognizability and difficulty was well-struck.

**Supplementary Table 2**

*Answer frequency for task-related follow-up questions*

| Question | Answer | | | |
|---|---|---|---|---|
| | Yes, very much. | Yes, mostly. | No, not very. | No, not at all. |
| Were the instructions clear and understandable? | 100% | 0% | 0% | 0% |
| Did you put a good effort into giving accurate answers? | 85.26% | 14.74% | 0% | 0% |
| Did you find the different objects distinguishable? | 6.32% | 47.37% | 41.05% | 5.26% |

Note: Response levels represent an approximation of the multiple-choice answer. For the exact phrasing, see Table 1.

Some participants reported using a strategy of trying to remember only Gain stimuli (N = 6) or only Loss stimuli (N = 3), which may have distorted their learning asymmetry. However, one-sample t-tests (two-sided) did not show that learning asymmetry for either subset differed significantly from zero, $t(5) = -0.97$, $p = .376$, $t(2) = 0.12$, $p = .917$, and so none of these cases were excluded. All other strategies that participants reported did not have a risk of distorting the results.

**Learning asymmetry and self-report measures**

**Supplementary Table 3**

*Descriptive statistics for self-report measures*

| Self-report scale | *M* | *SD* | Range |
|---|---|---|---|
| DASS – Depression | 4.71 | 3.99 | 16 |
| DASS – Anxiety | 4.39 | 3.44 | 18 |
| DASS – Stress | 7.13 | 3.68 | 16 |
| BAS – Drive | 2.59 | 0.53 | 2.75 |
| BAS – Reward Responsiveness | 3.39 | 0.36 | 1.60 |
| BAS – Fun Seeking | 2.88 | 0.49 | 2.25 |
| BIS | 3.20 | 0.44 | 2.17 |
| BSI – GSI | 0.82 | 0.57 | 2.74 |
| ASSIST | 22.57 | 19.81 | 99 |
| UPPS-P | 2.77 | 0.34 | 1.71 |

As an additional check on the impact of exclusion, two linear regression analyses were performed with the entire sample using the wAcc asymmetry as the dependent variable. All predictors combined did not explain a significant amount of variance, $F(6, 94) = 2.11$, $p = .059$, $R^2 = .119$. All predictors other than the PFS did not meet the inclusion criterion and were excluded. The final model consisted of only the PFS, $F(1, 99) = 6.14$, $p = .015$, $R^2 = .058$. For the second regression analysis, the General Severity Index (GSI) of the BSI, the sum score of the ASSIST, and the three subscales of the DASS were used as predictors, as they measure symptoms relevant to mental disorders. All predictors combined did not explain a significant amount of the variance, $F(5, 95) = .59$, $p = .706$, $R^2 = .030$. No predictors met the inclusion criterion, thus no model explained a significant amount of variance. Additional model statistics can be found in Supplementary Table 4.

**Supplementary Table 4**

*Summary of backwards entry regression analysis for personality traits and mental disorder symptoms, including the full sample.*

| Model | Predictors | *B* | SE *B* | *β* | SE *β* | *t* | *p* | 95% CI for B | | Tolerance |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper | |
| 1 | UPPS-P | 10.85 | 8.46 | .16 | .13 | 1.28 | .203 | -6.0 | 27.7 | .59 |
| | **PFS** | **7.29** | **3.24** | **.24** | **.11** | **2.25** | **.027** | 0.9 | 13.7 | .82 |
| | BIS | 10.77 | 5.42 | .22 | .11 | 1.99 | .050 | 0.0 | 21.5 | .79 |
| | BAS-D | 1.02 | 4.72 | .02 | .10 | 0.22 | .830 | -8.4 | 10.4 | .86 |
| | BAS-F | 6.49 | 6.16 | .13 | .13 | 1.05 | .295 | -5.7 | 18.7 | .58 |
| | BAS-R | -12.16 | 6.98 | -.18 | .11 | -1.74 | .085 | -26.0 | 1.7 | .84 |
| 1′ | **PFS** | **7.34** | **2.96** | **.24** | **.10** | **2.48** | **.015** | 1.5 | 13.2 | 1 |
| 2 | BSI-GSI | -2.58 | 8.33 | -.06 | .20 | -0.31 | .758 | -19.1 | 14.0 | .25 |
| | ASSIST | -0.02 | 0.15 | -.01 | .12 | -0.11 | .913 | -0.3 | 0.3 | .70 |
| | DASS-Depression | -1.36 | 1.23 | -.23 | .21 | -1.11 | .272 | -3.8 | 1.1 | .23 |
| | DASS-Anxiety | 1.07 | 1.06 | .16 | .16 | 1.01 | .316 | -1.0 | 3.2 | .40 |
| | DASS-Stress | 0.71 | 1.13 | .11 | .18 | 0.63 | .530 | -1.5 | 3.0 | .31 |
| 2′ | – | | | | | | | | | |

Note: Model 1 = personality traits, model 2 = mental disorder symptoms, prime (′) = final model. Coefficients significant at the $p < .05$ level are bolded.

**Supplementary Table 5**

*Bivariate correlations between learning asymmetry and self-report measures*

| | Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | wAcc asymmetry | — | | | | | | | | | | | | |
| 2 | Gain trial wAcc | .619** | — | | | | | | | | | | | |
| 3 | Loss trial wAcc | -.438** | .431** | — | | | | | | | | | | |
| 4 | DASS - Depression | -.117 | .006 | .139 | — | | | | | | | | | |
| 5 | DASS - Anxiety | .076 | .083 | .019 | .671** | — | | | | | | | | |
| 6 | DASS - Stress | -.008 | -.005 | .000 | .755** | .684** | — | | | | | | | |
| 7 | BAS – Drive | -.123 | -.125 | .005 | -.018 | -.103 | .010 | — | | | | | | |
| 8 | BAS – Reward Responsiveness | -.079 | -.039 | .058 | .099 | .032 | .199 | .223* | — | | | | | |
| 9 | BAS – Fun Seeking | -.098 | .031 | .146 | .132 | -.053 | .055 | .258* | .304** | — | | | | |
| 10 | BIS | .207* | .130 | -.108 | .399** | .516** | .506** | -.141 | .172 | -.087 | — | | | |
| 11 | BSI – GSI | -.064 | .030 | .105 | .824** | .715** | .741** | -.074 | .119 | .026 | .526** | — | | |
| 12 | ASSIST | -.144 | -.146 | -.018 | .515** | .284** | .491** | .003 | -.015 | .277** | .157 | .422** | — | |
| 13 | UPPS-P | .145 | .037 | -.114 | -.563** | -.322** | -.503** | -.134 | -.157 | -.563** | -.178 | -.447** | -.635** | — |
| 14 | PFS | .249* | .259* | .010 | .224* | .314** | .237* | -.172 | .148 | -.041 | .272** | .256* | .014 | -.175 |

Note: * $p < .05$. ** $p < .01$.

# Chapter 5

Learning Asymmetry as a Predictor of Mood and Behavior Dynamics: A Network Analysis

Submitted as: Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. Learning Asymmetry as a Predictor of Mood and Behavior Dynamics: A Network Analysis. *Journal of Behavior Therapy and Experimental Psychiatry.*



.

**Abstract**

While studying appetitive and aversive conditioning is common in psychopathology research, studies that measure both types of learning simultaneously are rare. To gain insight into the role of appetitive and aversive learning in the complex interaction of positive mood, negative mood, worry, craving, avoidance and impulsive behavior, this study used a relative measure of the strength of appetitive versus aversive learning – the learning asymmetry – as a predictor of network dynamics of mood states and behavior. 100 healthy volunteers performed an appetitive and aversive conditioning task and completed an ecological momentary assessment study, where they were surveyed six times per day for 21 days. Groups were defined based on higher sensitivity to appetitive learning (positive learning asymmetry) or aversive learning (negative learning asymmetry). The positive asymmetry group was hypothesized to be more sensitive to positive mood changes, and the negative asymmetry group was hypothesized to be more sensitive to negative mood changes. Results show that impulsive behavior was more likely to follow negative mood, specifically anger, in the positive but not the negative asymmetry group. These results demonstrate the potential for network analysis to elucidate complex interactions between mood and behavior associated with individual differences in learning.

## Introduction

Dealing with the unpredictable challenges in life is a universal struggle, and our mental health depends on how well we cope with these challenges. Individual differences play a role in many aspects of mental health, including sensitivity to reward and punishment. While people generally learn from experience to seek out pleasant, rewarding outcomes and avoid unpleasant, aversive outcomes, individual variations in learning can lead to maladaptive behaviors. For example, some may develop excessive avoidance due to oversensitivity to aversive outcomes while others may engage in overly risky behavior due to heightened sensitivity to appetitive outcomes. Such an imbalance between appetitive and aversive learning may play a role in the development of psychopathology including addictive disorders, mood disorders and anxiety disorders. This highlights how important individual differences are in appetitive and aversive learning – which are generally studied through conditioning paradigms – and how they can impact mental health.

In conditioning studies, participants are presented with a neutral stimulus (conditioned stimulus or CS) that is paired with a salient stimulus (unconditioned stimulus or US). Participants learn to expect the US after the CS, which is reflected by, for example, their reported expectancy or heart rate. Such responses are compared with a CS that has not been paired with a US (known as the CS-, in contrast with the CS+) to analyze the effect of the CS-US association. A long-standing area of research explores how fear and anxiety disorder originate and persist, through mechanisms such as heightened responsiveness to aversive stimuli (Pitman & Orr, 1986). Anxiety disorder patients show increased fear responses to the CS- compared to healthy controls, according to meta-analyses (Duits et al., 2015; Lissek et al., 2005), suggesting reduced safety learning. Aversive learning has also been studied in the context of other mental disorders: men at risk for alcoholism are less able to discriminate between aversive and neutral stimuli, but this is driven by reduced responding to the aversive stimulus, rather than to the neutral stimulus (Finn et al., 1994). These studies show that even within aversive conditioning, CS+/CS- learning shows different effects in different study populations.

While the research on aversive learning and anxiety is extensive, research into appetitive learning and its relation to psychopathology is comparatively uncommon and mostly focuses on obesity and eating disorder research. Although obesity is not a mental disorder, some studies suggest that disordered appetitive conditioning plays a role in both obesity (van den Akker et al., 2018) and eating disorders like bulimia nervosa (Jansen et al., 1992). In both cases, strong appetitive CS-US associations are hypothesized to drive overeating, but this is not consistently supported: some studies demonstrate that overweight/obese individuals show poor discrimination of appetitive and neutral stimuli relative to those with healthy weight (Coppin et al., 2014; van den Akker et al., 2017; Zhang et al., 2014), yet others find the reverse, namely better CS+/CS- discrimination (Meemken et al., 2018; Meyer et al., 2015) or report no learning differences depending on weight (van den Akker et al., 2019).

As the previous studies show, there are significant gaps in our understanding of appetitive and aversive learning and their relation to psychopathology, and this may be due to the rarity of studies that investigate appetitive and aversive learning simultaneously. In cases where the US is particularly unpleasant, the relief of receiving no US is comparatively pleasant, and when the participant expects a particularly rewarding US, its omission is comparatively unpleasant. Thus, it has been argued that the omission of a US may also be perceived as appetitive or aversive (Gray, 1975). As a result, studies that measure learning using on only appetitive or aversive stimuli may lack the proper context for their results. For example, the reduced safety learning found in anxiety disorder studies may in fact be an expression of reduced appetitive learning. To properly account for this, a measure was used that incorporated appetitive and aversive learning, namely: learning asymmetry (Shook et al., 2007; anonymized reference A). This was measured using a task in which participants were presented with both appetitive and aversive USs, and their learning asymmetry was determined by which type of association they learned more effectively. Specifically, learning asymmetry was calculated as the difference between the strength of appetitive and aversive associations, creating a relative measure of learning that indicates how sensitive the participant is to the type of stimulus.

Two studies investigated learning asymmetry for its association with personality traits and psychopathology symptoms. Using a system of point gain and loss as appetitive and aversive stimuli, one study examined the relation of impulsivity, neuroticism, and anhedonia with learning asymmetry (anonymized reference A).

Only impulsivity was shown to be associated with learning asymmetry: More impulsive individuals showed weaker aversive learning, but showed no difference in appetitive learning. In another study, learning asymmetry was measured using sweet and bitter tastes to examine its relation to psychological distress and substance use, in addition to impulsivity and anhedonia (anonymized reference B). Psychological distress was associated with weaker appetitive learning for participants with high psychological distress; aversive learning showed no difference. Substance use, impulsivity, and anhedonia were not associated with learning asymmetry. Together, these studies suggest that learning asymmetry may be selectively associated with certain psychological traits, with impulsivity linked to reduced aversive learning and psychological distress to diminished appetitive learning.

The current study builds on previous results using network analysis of ecological momentary assessment (EMA) data. EMA captures real-time data on symptoms and behaviors in participants' daily lives, repeatedly throughout the day, using devices such as smartphones. Data collection via EMA thus reduces recall bias, improves ecological validity, and allows for examining variables in higher detail. This approach allows us to determine whether there are mood and behavioral dynamics that are characteristic of positive or negative learning asymmetry individuals and so improve our understanding of how individual differences in learning may lead to maladaptive behavior. Specifically, by measuring how often participants report certain moods and behaviors in their daily lives, we can explore whether individuals with a positive or negative learning asymmetry differ in their sensitivity to positive or negative moods. Additionally, we can examine whether and how moods are linked to states like worrying and craving, or behaviors such as impulsivity or avoidance, and how these relations differ between positive and negative learning asymmetry.

To this end, we measured learning asymmetry as in (anonymized reference A), namely with a conditioning task using abstract 3D objects as conditioned stimuli and a system of point gain and loss, tied to an additional monetary reward, as unconditioned stimuli. Then, participants performed EMA for a period of three weeks and were asked about their mood, craving, impulsive behavior, and avoidance behavior six times per day. Regression analysis and time series analysis were used to compare the participants with positive and negative learning asymmetry. Time series analysis examines how measurements at individual time points affect each other. By applying time series analysis to EMA data, we can discover how changes in one variable are followed by changes in other variables on the next time point. In this case, we investigated whether participants in the positive or negative learning asymmetry groups responded differently to the EMA questionnaires overall, and whether the pattern of responses from one time point to the next differed between groups.

We hypothesized that differences in learning asymmetry predict differences in average levels of reported behavior, in addition to time series effects of mood and behavior interactions. Our primary hypothesis, regarding group averages, is that the positive asymmetry group would show increased impulsive action and craving, as a positive learning asymmetry was previously shown to be associated with impulsivity. Conversely, we hypothesize that the negative asymmetry group shows increased avoidance, as we expect this group is more sensitive to aversive learning. Our secondary hypothesis, regarding time series effects, is that impulsive behavior, avoidance and craving would show differences in time-series effects between groups. Given that learning asymmetry is a behavioral measure, it seems most likely that the network will show differences in impulsive behavior and avoidance between groups. We expect that the positive asymmetry group, which may be more impulsive and less sensitive to aversive learning, would be more sensitive to these behaviors causing positive mood, due to their higher sensitivity to appetitive reinforcement. Conversely, we expect that the negative asymmetry group, which may be less sensitive to appetitive learning, would be more sensitive to these behaviors causing negative mood, due to their higher sensitivity to negative reinforcement.

## Method

### Participants

101 undergraduate students participated in the study. They were recruited through the university's research participation system for psychology students. Participants were between 18 and 26 years of age, $M = 20.17$, $SD = 1.63$, and 89% was female. Participants' national origin was predominantly German (42.6%) and

Dutch (24.8%) with other European nationalities accounting for 22.8% of the sample and non-European nationalities for 8.9%. Participation occurred between 26 January 2024 and 30 April 2024. Participants took part in the study for either study credit or financial compensation.

## Materials

### *Learning Task Stimuli*

Stimuli were generated using the scripts and instructions provided by Watson et al. (2019) on 3DS Max 2022. Twenty-six stimuli were used, ensuring that no two stimuli shared more than one feature between them. Thirteen stimuli were associated with point gain and thirteen with point loss. These associations were counterbalanced between participants, and assigned such that no single feature was consistently associated with gain or loss.

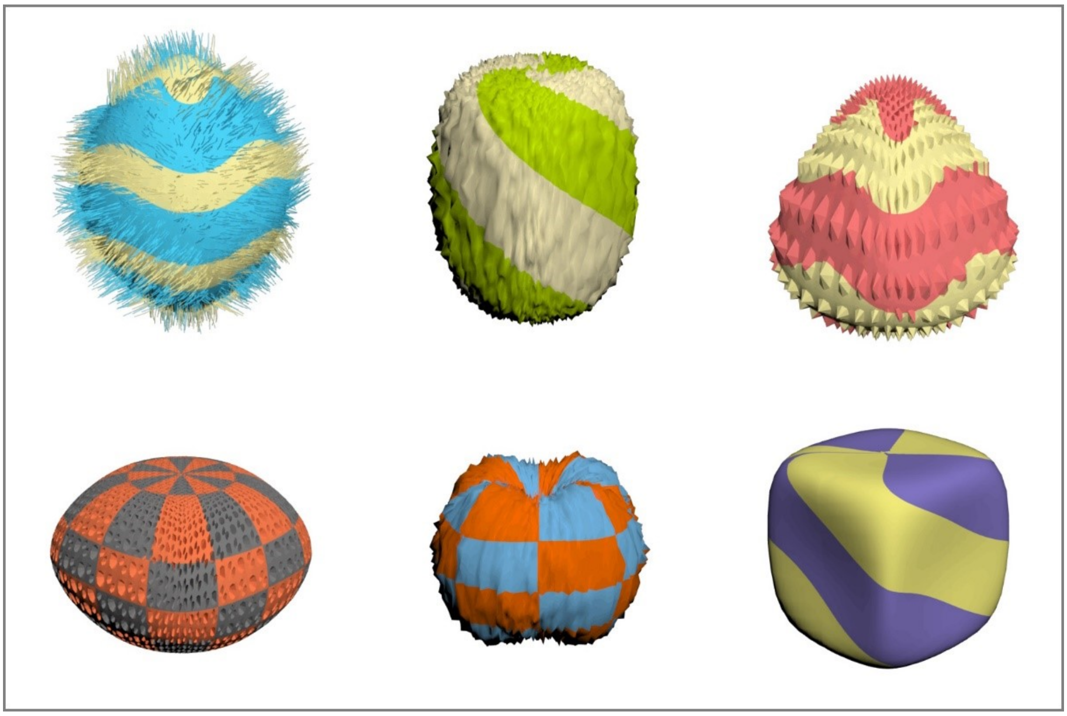


*Figure 1: Six examples of Quaddles used as conditioned stimuli.*

### *Task Performance*

After completing the task, participants were asked nine questions about various aspects of the task, to verify that they completed the task as intended and that no problems occurred with the online test platform. Details on these questions are available in Supplementary Table 1.

### *EMA Questionnaires*

Participants answered questionnaire items on a 7-point Likert scale. The top option was labeled "Not at all" and the bottom option "Extremely". The remaining options were numbered 2 through 6. The momentary questionnaires contained 23 questions (see Table 1). Daily questionnaires consisted of a single yes/no question, namely "Was today extremely out of the ordinary for you for any reason? For example, because of significant life events like a death in the family, or public holiday celebrations like carnaval. Please answer 'Yes' only if the events affected the entire day." This question was used to determine whether days with unrepresentative data should be excluded.

### *Avicenna*

Avicenna Research (formerly Ethica Data) provides software specialized for data collection using personal mobile devices. The Avicenna Android and iOS apps were used to deliver notifications to participants and allow them to fill out momentary, daily and evaluation questionnaires. Momentary questionnaires were delivered six times per day for 21 days at irregular moments within two-hour time windows, distributed normally around the center of the time window. The time windows at which measurements started and ended each day were adapted to the usual time that participants reported to be awake, for which they could choose 8:00-22:00, 10:00-0:00, or 12:00-2:00. All participants received questionnaires within two-hour time windows starting at 13:00, 15:00, 17:00, and 19:00. Depending on participants' awake time, earlier questionnaires were delivered in the time windows starting 9:00 and 11:00, and later questionnaires starting 21:00 and 23:00. Daily questionnaires were delivered at the end of the latest time window for momentary questionnaires.

When a momentary questionnaire became available, participants received one notification immediately, and another notification after 20 minutes. The questionnaire expired if they did not answer within 30 minutes of the first notification. Daily questionnaires and the evaluation questionnaire did not expire.

## Procedure

### *Learning asymmetry task*

Ethical approval was obtained from the local Ethics Review Committee Psychology and Neuroscience (approval code OZL_234_27_02_2021_S7_A1). Participants were first invited to the lab and seated at a computer in a soundproofed room, where they were provided with an information letter describing the task and the EMA protocol, and signed the informed consent form. They then received verbal instructions for the learning task, including that they would receive an additional reward (a €5 voucher) if they scored more than a predetermined number of points. The first block (the acquisition block) showed the participant the association of each of the twenty stimuli, and no responses were necessary. The subsequent three blocks (the response blocks) repeated the same stimuli and asked the participant to respond to each stimulus based on the associations they learned from the acquisition block.

The acquisition block indicated the point value of each stimulus ("Gains 10 points"/"Loses 10 points"). During the subsequent three response blocks the participant was asked to 'accept' or 'reject' the outcome of each stimulus. The goal was to gain points and avoid losing points by 'accepting' stimuli associated with point gain and 'rejecting' stimuli associated with point loss. This was done by responding to the question "Accept the result?" by clicking one of two buttons, 'Yes' or 'No'. On trials where the stimulus signaled point gain (hereafter: 'Gain trials'), responding 'Yes' would add 10 points. Conversely, on trials where the stimulus signaled point loss (hereafter: 'Loss trials'), responding 'Yes' would subtract 10 points. Responding 'No' resulted in neither point gain nor loss. In this way, Gain trials created appetitive associations through the possibility of gaining points, and Loss trials created aversive associations through the possibility of losing points.

After responding to each trial, participants selected a point on a 100 point visual analog scale to indicate their confidence in their decision. They then received feedback indicating the value of the stimulus. In addition, the participant's point total for the entire task was displayed with each instance of feedback. For more details, see Figure 2.

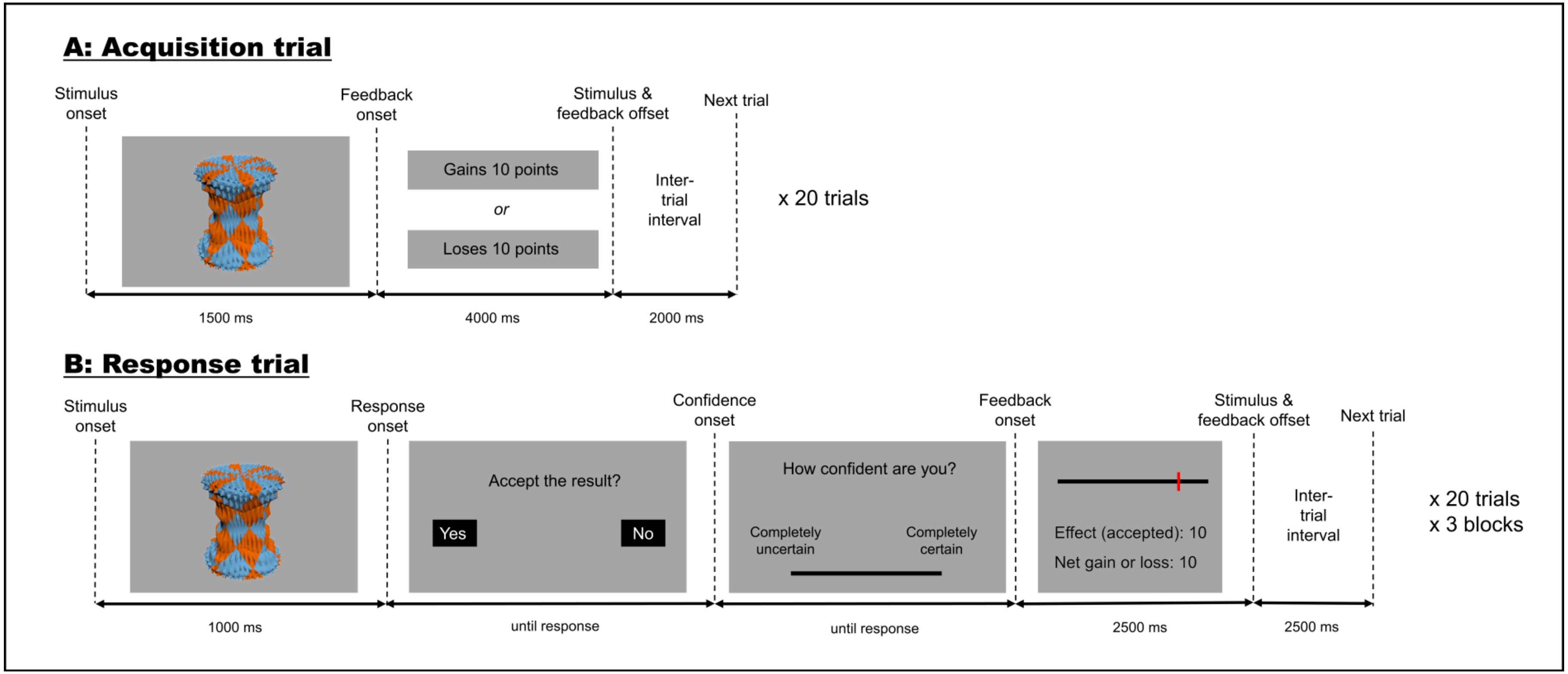



*Figure 2: Diagram of the task procedure for the learning task, split by Acquisition trials (A) and Response trials (B). Areas in grey represent elements displayed during the task, but do not represent the entire screen. The feedback shown for the response trial is an example for a Yes response to a Gain stimulus.*

## Outcome measures

From the measures of accuracy and confidence, a composite variable was created using the sum of confidence ratings on all correct trials minus the sum of confidence ratings on all incorrect trials, then divided by the total number of trials, a value that will henceforth be referred to as 'weighted accuracy' or 'wAcc'. By weighting participants' answers according to their confidence, wAcc more closely represents the strength of the learned associations.

Learning asymmetry was calculated by averaging wAcc for both Gain and Loss trials across all three blocks, then subtracting the wAcc for Loss trials from the wAcc from Gain trials. Participants were then divided into groups depending on whether their learning asymmetry was positive or negative.

From the EMA data, items (see Table 1) were grouped into the nodes Positive Mood (1, 6-8), Negative Mood (2-5), Craving (9-13), Impulsive Behavior (15-19) and Avoidance (20-23) by averaging the responses belonging to each node at each time point. Worry (14) was also included as a node consisting of one item. Grand averages of these nodes were obtained by averaging all time points together for each participant.

## Transparency and Openness

The study's hypotheses, analyses, exclusion criteria and sample size were preregistered at https://aspredicted.org/3w3d-p9ch.pdf. Stimulus materials and task script are available at https://osf.io/pdn6w/?view_only=5830913ecb114cbbb78e77f6e289826e. Data is available upon reasonable request.

**Table 1**

*Items from the EMA questionnaire*

| Question |
|---|
| Mood |
| 1. At this moment, how happy do you feel? |
| 2. At this moment, how sad do you feel? |
| 3. At this moment, how anxious do you feel? |
| 4. At this moment, how stressed do you feel? |
| 5. At this moment, how angry do you feel? |
| 6. At this moment, how calm do you feel? |
| 7. At this moment, how energetic do you feel? |
| 8. At this moment, how content do you feel? |
| Craving |
| 9. Do you currently have a craving for snacks or junk food? |
| 10. Do you currently have a craving for alcohol? |
| 11. Do you currently have a craving for smoking or vaping? |
| 12. Do you currently have a craving for another drug? |
| 13. Do you currently have a craving for something else not mentioned? |
| Worry |
| 14. Since the last beep, I have been worrying. |
| Impulsive behavior |
| 15. Since the last beep, I lost control when eating. |
| 16. Since the last beep, I said things without thinking. |
| 17. Since the last beep, I spent more money than I meant to. |
| 18. Since the last beep, I have felt impatient. |
| 19. Since the last beep, I have made a 'spur of the moment' decision. |
| Avoidance |
| 20. Since the last beep, I have avoided one or more difficult or awkward situations. |
| 21. Since the last beep, I have avoided one or more people. |
| 22. Since the last beep, I did something to prevent a bad thing from happening. |
| 23. Since the last beep, I tried to prevent a bad thing from happening by not doing something. |

*Note: Impulsive behavior items adapted from Tomko et al. (2014).*

## Results

### Inclusion and Exclusion

One hundred participants completed the entire EMA data collection period. Upon evaluating response times, it was found that one participant took less than fifteen seconds on average to complete the 23-item questionnaire, which was deemed insufficient time to give well-considered answers and this participant was excluded. The analyses were therefore performed on the remaining 99 participants.

### Task Performance

To determine whether participants learned from Gain and Loss trials, a repeated-measures analysis of variance (ANOVA) was performed using Valence (Gain vs. Loss) and Block (1 vs. 2 vs. 3) as factors and wAcc as the dependent variable. No effect of Valence was found, $F(1, 98) = 2.18$, $p = .143$, but the effect of Block, $F(2, 97) = 52.30$, $p < .001$*, and the interaction between Block and Valence, $F(2, 97) = 8.35$, $p < .001$*, were significant. Paired t-tests were then performed for each pair of consecutive blocks of the same valence and their wAcc. Significance thresholds for each test were determined by the Holm-Bonferroni method. The difference between wAcc on Gain and Loss trials was not significant for block 1, $t(98) = -2.18$, $p = .031$, or block 2, $t(98) = 1.69$, $p = .098$, but it was significant for block 3, $t(98) = 3.01$, $p = .003$*. Loss trials showed a higher wAcc than Gain trials in block 3. Furthermore, the difference in wAcc for Gain trials between block 1 and 2 was not significant, $t(98) = 1.94$, $p = .056$, but it was significant between block 2 and 3, $t(98) = 4.39$, $p < .001$*. For loss trials, the difference between block 1 and 2 was significant, $t(98) = -3.71$, $p < .001$*, and also between block 2 and 3, $t(98) = -7.31$, $p < .001$*, with wAcc increasing in every block for loss trials, and between block 2 and 3 in gain trials. The interaction effect was therefore driven by wAcc on Loss trials increasing more than wAcc on Gain trials. Although wAcc on Gain trials did not significantly increase between block 1 and 2, wAcc for both Gain and Loss trials significantly increased by the final block of the task, indicating that participants learned the associations better over time.

### *Learning Asymmetry*

Learning asymmetry did not differ significantly from zero, $t(98) = -1.30$, $p = .198$. 51 participants showed a positive learning asymmetry, and 48 participants showed a negative learning asymmetry. Therefore, group comparisons could proceed without need for adjusting group allocation.

## Ecological Momentary Assessment

### *Compliance*

Compliance was satisfactory, with the average of answered questionnaires exceeding 80% (see Table 2).

**Table 2**

*Descriptive statistics for compliance and missing streaks*

| Variable | *M* | *SD* | Minimum | Maximum |
|---|---|---|---|---|
| Compliance (%) | 83.08 | 12.42 | 47.62 | 100 |
| Missing streaks | 1.35 | 0.54 | 0 | 4.63 |

*Note: Compliance: percentage of momentary questionnaires finished over the entire 21-day measurement period, missing streaks: average number of consecutively unanswered questionnaires.*

### *Grand averages*

One-sample t-tests were performed on each EMA node average to determine whether they differed significantly from 1 ("Not at all"). All node averages were significantly higher than 1, all *p's* < .001. Test statistics can be found in Table 3.



.

**Table 3**

*One-sample t-test statistics for EMA nodes*

| Node | $t$ | $df$ | $p$ | Mean Difference | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|---|---|
| Positive mood | 40.49 | 98 | <.001 | 2.68 | 2.55 | 2.81 |
| Negative mood | 16.00 | 98 | <.001 | 1.22 | 1.07 | 1.37 |
| Worry | 15.94 | 98 | <.001 | 1.69 | 1.48 | 1.90 |
| Craving | 10.21 | 98 | <.001 | 0.42 | 0.34 | 0.50 |
| Impulsive behavior | 12.15 | 98 | <.001 | 0.66 | 0.55 | 0.77 |
| Avoidance | 7.92 | 98 | <.001 | 0.46 | 0.35 | 0.58 |

*Note: CI: Confidence Interval.*

To test the primary hypothesis, an average of all time points of each EMA node was calculated for each participant and tested against learning asymmetry to determine whether they showed overall higher or lower item responses irrespective of the time of measurement. Two regression analyses were performed using learning asymmetry as the dependent variable. In the first analysis, learning asymmetry was used as a continuous predictor, and in the second analysis, positive and negative learning asymmetry groups were used as a categorical predictor. This procedure allows for the comparison of these results with group-based analysis of time-series effects, which is necessary as time-series effects may be influenced by differences in response averages. The model using continuous learning asymmetry as the dependent did not explain a significant amount of variance, $F(6, 92) = 0.5$, $p = .807$, $R^2 = 0.032$, nor did the model using learning asymmetry groups, $F(6, 92) = 0.7$, $p = .638$, $R^2 = 0.045$. Further model statistics can be found in Supplementary Table 4. These data show that participants did not consistently report higher or lower levels of their mood or behavior depending on their learning asymmetry.

*Time-series analysis*

Time-series data were analyzed using multilevel vector autoregressive (VAR) models as per Haslbeck et al. (2023). VAR models estimate temporal dynamics by modeling each variable as a linear combination of all variables, including that same variable at previous time points. By fitting VAR models for individual participants' EMA data, a multilevel VAR model distinguishes variance caused by between- and within-participant effects. We compare multilevel VAR models between positive and negative learning asymmetry groups using a permutation test, which creates 1000 random allocations of participants to two groups and tests the differences in standardized beta coefficients found between learning asymmetry groups against those found between random groups.

The validity of the model was examined by plotting the residuals and $R^2$ values of individual VAR models per node. These can be found on pp. 133. The full networks are shown in Figure 3. The strongest inter-node standardized beta in the networks was present in both groups, namely between negative mood and worry. Strong autocorrelations, that is, correlations between the response on a node and the response on that node on the next time point, were also present in both groups for positive mood and negative mood, and to a lesser extent, craving and worry. Significant differences between the groups were limited to the nodes for negative mood, impulsive behavior, and avoidance. The positive asymmetry group showed a greater autocorrelation for avoidance, while the negative asymmetry group showed a greater autocorrelation for impulsive behavior. Furthermore, the positive asymmetry group showed more impulsive behavior following negative mood, an effect that was absent in the negative asymmetry group. As the analysis on pp. 123 did not find any significant differences in node averages, we can be confident that time-invariant response differences between groups, such as an overall higher rate of avoidance responses for the positive asymmetry group, did not affect these results.

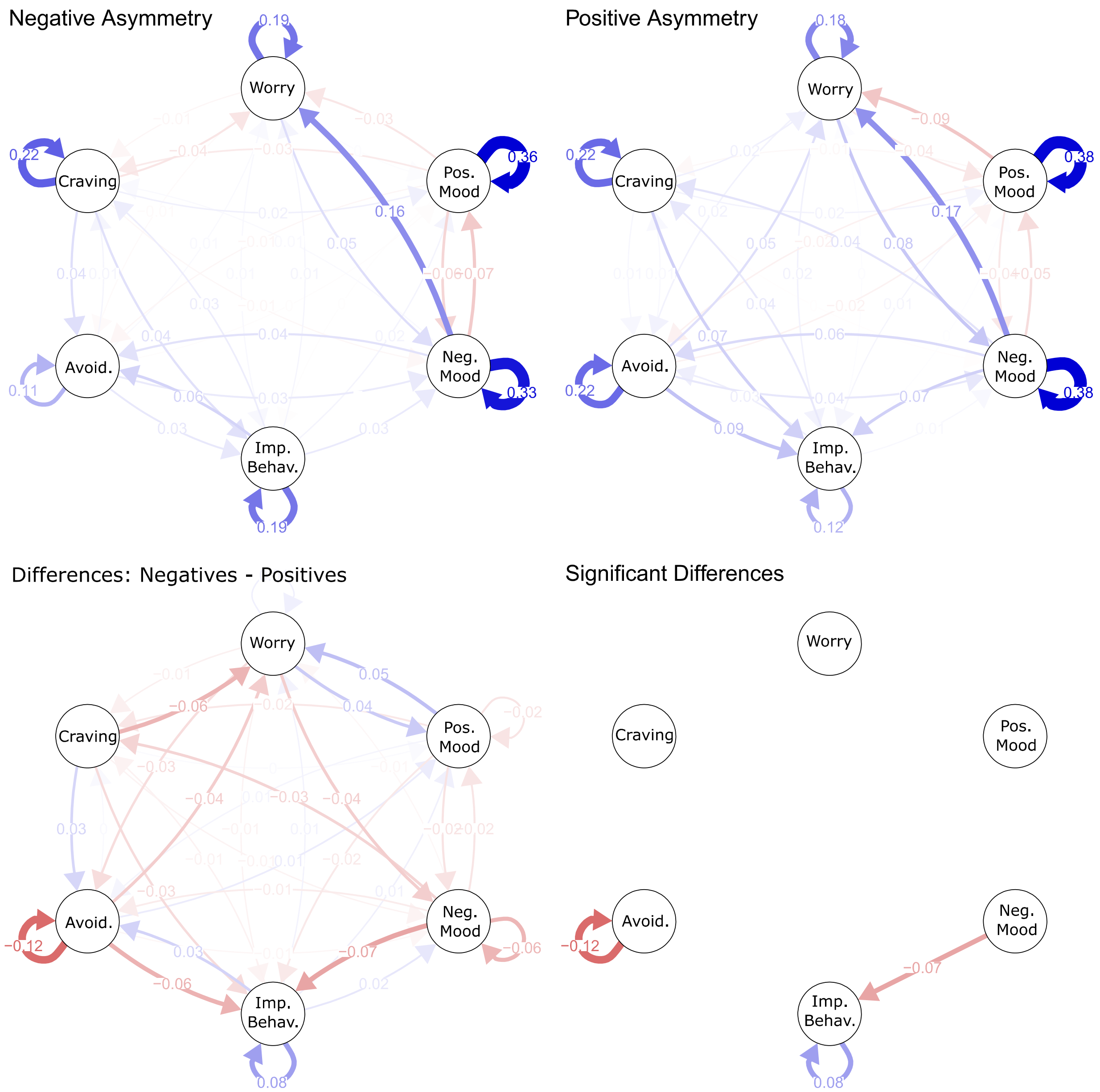



*Figure 3: Network characteristics of EMA nodes. Arrows show differences in standardized beta coefficients between a node at time point t – 1 and time point t.*

As the individual EMA items were aggregated into nodes to improve statistical power, further exploratory analyses could be performed on the disaggregated items to further investigate these effects. The permutation analysis was repeated using the anger, sadness, anxiety and stress EMA items, as well as the impulsive behavior node. The networks from this analysis are visualized in Figure 4. Anger was significantly more often followed by sadness, anxiety, and impulsive behavior in the positive than the negative asymmetry group, suggesting that the effect of negative mood in the previous network was driven primarily by anger.

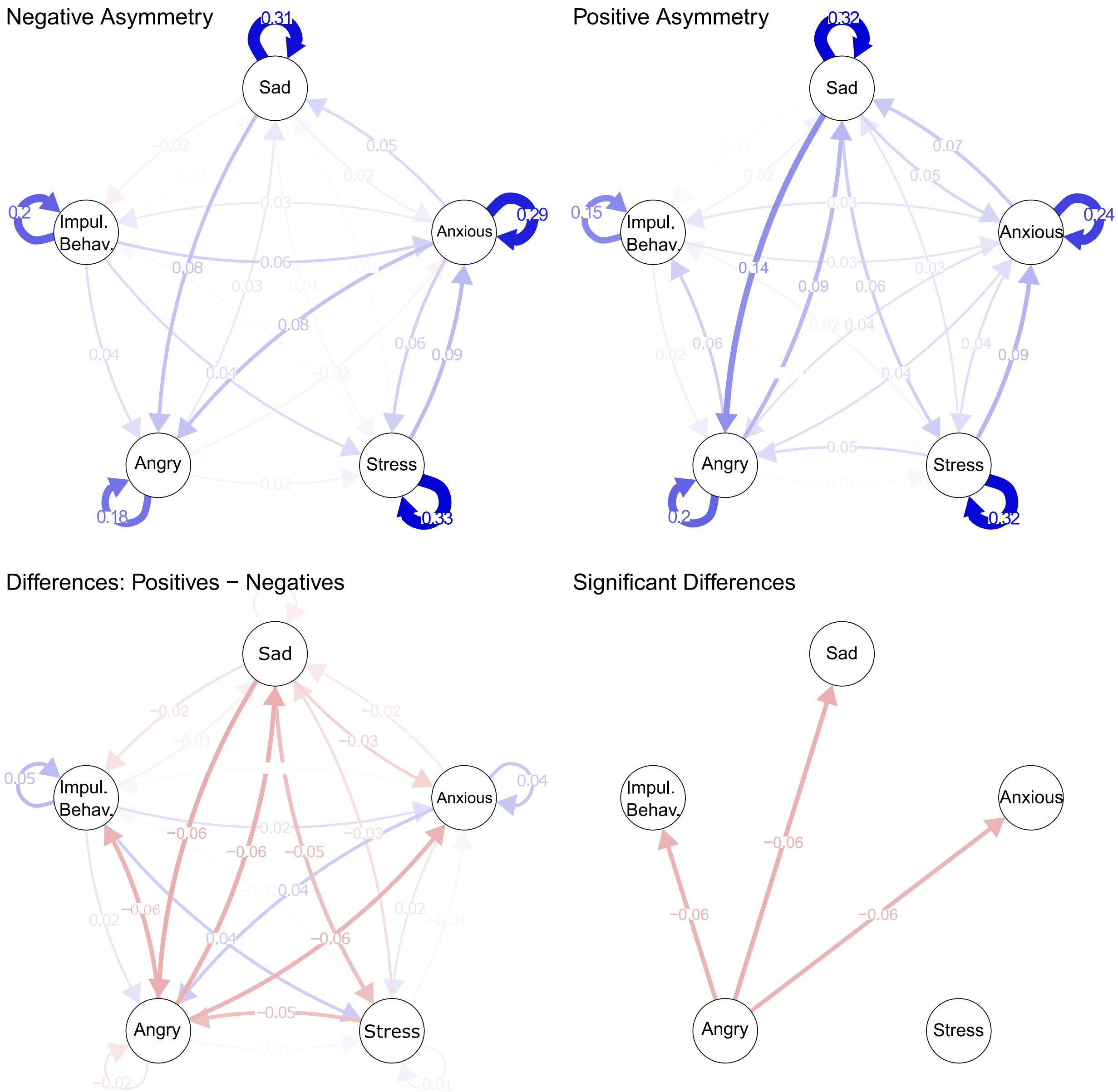



*Figure 4: Network characteristics of negative mood items and impulsive behavior. Arrows show standardized beta coefficients between a node at time point t – 1 and time point t.*

## Discussion

In the positive asymmetry group, negative mood was positively correlated with impulsive behavior, while this effect was absent in the negative asymmetry group. Disaggregating the negative mood items showed that anger was the only item that significantly correlated with impulsive behavior, indicating that anger was followed by impulsive behavior. In addition, avoidance was more strongly autocorrelated in the positive than the negative asymmetry group, while the reverse was true for the autocorrelation of impulsive behavior. These autocorrelations do not imply a greater or lesser degree of avoidance or impulsive behavior: instead, they indicate the degree to which participants' responses on this node are consistent with their previous responses. While autocorrelations on their own are difficult to interpret, a high autocorrelation may indicate that the predictive value of all other nodes combined is lower. Therefore, the finding that the negative asymmetry group shows a lower avoidance autocorrelation than the positive asymmetry group indicates that, within the negative asymmetry group, all other nodes played a larger role in predicting avoidance, even if no singular node was significantly different in its association with avoidance. Similarly, the positive asymmetry group showed a lower impulsive behavior autocorrelation than the negative asymmetry group, indicating that all other nodes were more predictive of impulsive behavior than impulsive behavior on previous time points in the positive asymmetry group. While this does not give us any clue towards the behavioral dynamics that result in avoidance or impulsive behavior, it does indicate that positive and negative learning asymmetry individuals behave differently from each other in ways that may be elucidated by future research. Overall, these findings are consistent with the learning asymmetry concept. The negative asymmetry group, characterized by stronger aversive learning, was more reactive in terms of avoidance, and the positive asymmetry group, characterized by stronger appetitive learning, was more reactive in terms of impulsive behavior. In other words, participants' sensitivity to appetitive and aversive learning was apparent in their tendency towards avoidance or impulsive behavior in their everyday lives. While these effects were not apparent in the averaged responses, this is likely because the strength of EMA data lies in time-series analysis, and averaging all time points together results in a significant loss of detail and accuracy.

The positive asymmetry group was found to behave more impulsively following a higher negative mood. This effect has also been found in previous studies on impulsivity, namely in the trait of (negative) urgency, which is defined as a tendency to commit rash or regrettable actions as a result of intense negative affect (Whiteside & Lynam, 2001). It has also been studied using EMA, showing that self-reports of negative urgency are reflected by time-series effects between negative mood and impulsive behavior (Feil et al., 2020; Sharpe et al., 2021; Sperry et al., 2018). In a previous study measuring learning asymmetry with the same task, impulsivity was also associated with a positive learning asymmetry (anonymized reference A), as a result of weaker aversive learning rather than stronger appetitive learning among high impulsive individuals.

Conversely, the negative asymmetry group did not show inter-node effects that did not also occur in the positive asymmetry group, but it did show greater reactivity to avoidance. Previous study applying network analysis to avoidance have found that depressed mood predicts avoidance in anxiety disorder patients (Meine et al., 2024; Piccirillo & Rodebaugh, 2022) Given that, in a previous study, a negative learning asymmetry was associated with higher psychological distress (anonymized reference B), we expected to find that the negative asymmetry group would show higher negative mood in response to changes in other behavior, but this was not the case, nor did autocorrelations for negative mood differ between groups. It is possible that these effects were not found because the number of participants experiencing psychological distress was relatively small in the negative asymmetry group. By contrast, variation in impulsivity between participants may have been higher than that of psychological distress, making the effect of impulsive behavior easier to detect.

Combined, these results suggest that sensitivity for aversive stimuli may also play a role in how impulsivity leads to maladaptive behavior, emphasizing the importance of studying both appetitive and aversive learning simultaneously to test their effects on behavior. In addition, this shows that group comparisons of time series effects have potential for elucidating differences in complex behavior interactions between groups that differ on a certain trait.

There are several limitations pertaining to the methods of EMA and network analysis as applied in this study. First, the reliability metrics of the data and VAR models indicated some points of weakness. The overall response averages (see Table 3) and residual analyses (see pp. 136) show that some nodes have small deviations from 1, suggesting that participants were less likely to report impulsive behavior, avoidance, or craving than positive mood, negative mood and worry. This is expected to some degree in a sample of healthy volunteers who are not expected to show pathological avoidance or impulsivity, and while a greater degree of variation is more advantageous in terms of statistical power, there are challenges to achieving this in practice. Another network analysis study asks avoidance questions with a frequency of only one per day (Jover Martinez et al., 2025), but current network analysis methods do not allow comparing items with different assessment frequencies. Alternatively, the wording of the questions could be adjusted so that participants are more likely to report when they are feeling impulsive or avoidant, but this also increases the risk that the data no longer measures the behavior that we are interested in. Other EMA designs may consider using different response options for questions that concern the frequency of a behavior, given that the response options used were the same as those for reporting mood states. Alternatively, participants may be screened more strictly to filter out non-responders on certain measures of interest.

Second, the current analysis grouped participants by positive or negative learning asymmetry, or in other words, whether they were more sensitive to appetitive or aversive learning. This collapsed all variation in learning asymmetry into a binary variable, which may have resulted in a loss of sensitivity for certain effects that depend on how strong learning asymmetry is. Additionally, a learning asymmetry of zero was chosen as the split, but a large number of participants have a learning asymmetry close to zero, and they were categorized as positive or negative based on small differences between them, further reducing sensitivity. The current method of network analysis did not allow for time series analysis using a continuous variable as a moderator, but these methods are actively being developed, and future EMA studies may be able to make use of improved analysis methods.

In conclusion, the current results show that using EMA data to investigate group differences relevant to impulsive behavior and avoidance is a viable approach to study how individual differences affect complex emotional and behavioral dynamics. While the sample of healthy volunteers may have limited the effects we were able to discover, the association between anger and impulsive behavior is nevertheless consistent with previous research and was not detectable when examining the group averages. Future studies can gain more insight into the emotional and behavioral dynamics relevant to appetitive and aversive learning by examining populations of interest in further detail.

**Supplementary Material**

## Participant Task Evaluation

### Supplementary Table 1

*Questions and Answers from the Task Follow-Up Questionnaire*

| Question | Answers |
|---|---|
| 1. Were the tasks and instructions clear and understandable? | a) Yes, I understood everything perfectly<br>b) Yes, I understood most of it<br>c) No, there were a lot of things I didn't understand<br>d) No, I didn't understand any of it |
| 2. Please elaborate on what you found unclear. | |
| 3. Did you put a good effort into giving accurate answers? | a) Yes, I did the best that I could<br>b) Kind of, I didn't try as hard as I could have<br>c) Not really, I wasn't paying much attention<br>d) No, I was just clicking randomly |
| 4. Did you use any memorization techniques or other methods that helped you respond accurately? If yes, please describe them. If no, you can leave this blank. | |
| 5. Which of these words match your feelings about the task? Please select all that apply. | ▪ Interesting ▪ Boring<br>▪ Frustrating ▪ Difficult<br>▪ Challenging ▪ Rewarding |
| 6. How distinguishable did you find the different objects? | a) They were all clearly different |

| | |
|---|---|
| | b) They were somewhat similar but still distinguishable<br>c) They had a lot of similarities and I mixed some of them up<br>d) They were so similar I mixed them up constantly |
| 7. Are there any other thoughts or opinions you had about the tasks? Please list any that you can think of. | |

Note: Empty cells in the 'Answers' column indicate that participants were allowed to type in their own answer.

In the task-related follow-up questions, all participants indicated that the instructions were fully understandable and that they made at least a moderate effort to perform well (see Supplementary Table 2). When asked about the distinguishability of the stimuli, 52.53% of participants reported them to be mostly or very distinguishable, which is expected given that the stimuli consisted of many combinations of a few different properties. As such, these answers indicated that the balance between recognizability and difficulty was well-struck.

Some participants reported using a strategy of trying to remember only Gain stimuli (N = 6) or only Loss stimuli (N = 3), which may have distorted their learning asymmetry. However, one-sample t-tests (two-sided) did not show that learning asymmetry for either subset differed significantly from zero, $t(5) = -0.97$, $p = .376$, $t(2) = 0.12$, $p = .917$, and so none of these cases were excluded. All other strategies that participants reported did not have a risk of distorting the results.

**Supplementary Table 2**

*Answer frequency for task-related follow-up questions*

| Question | Answer | | | |
|---|---|---|---|---|
| | Yes, very much. | Yes, mostly. | No, not very. | No, not at all. |
| Were the instructions clear and understandable? | 100% | 0% | 0% | 0% |
| Did you put a good effort into giving accurate answers? | 85.86% | 14.14% | 0% | 0% |
| Did you find the different objects distinguishable? | 5.05% | 47.48% | 42.42% | 5.05% |

Note: Response levels represent an approximation of the multiple-choice answer. For the exact phrasing, see Table 1.

## Ecological Momentary Assessment

**Supplementary Table 3**

*Correlation coefficients between EMA node averages*

| | Positive Mood | Negative Mood | Craving | Impulsive Behavior | Avoidance |
|---|---|---|---|---|---|
| Positive Mood | — | | | | |
| Negative Mood | -.44 | — | | | |
| Craving | -.17 | .52 | — | | |
| Impulsive Behavior | -.09 | .63 | .57 | — | |
| Avoidance | -.15 | .61 | .47 | .82 | — |
| Worry | -.38 | .78 | .36 | .58 | .54 |

**Supplementary Table 4**

Summary of hierarchical regression analysis for EMA nodes predicting learning asymmetry

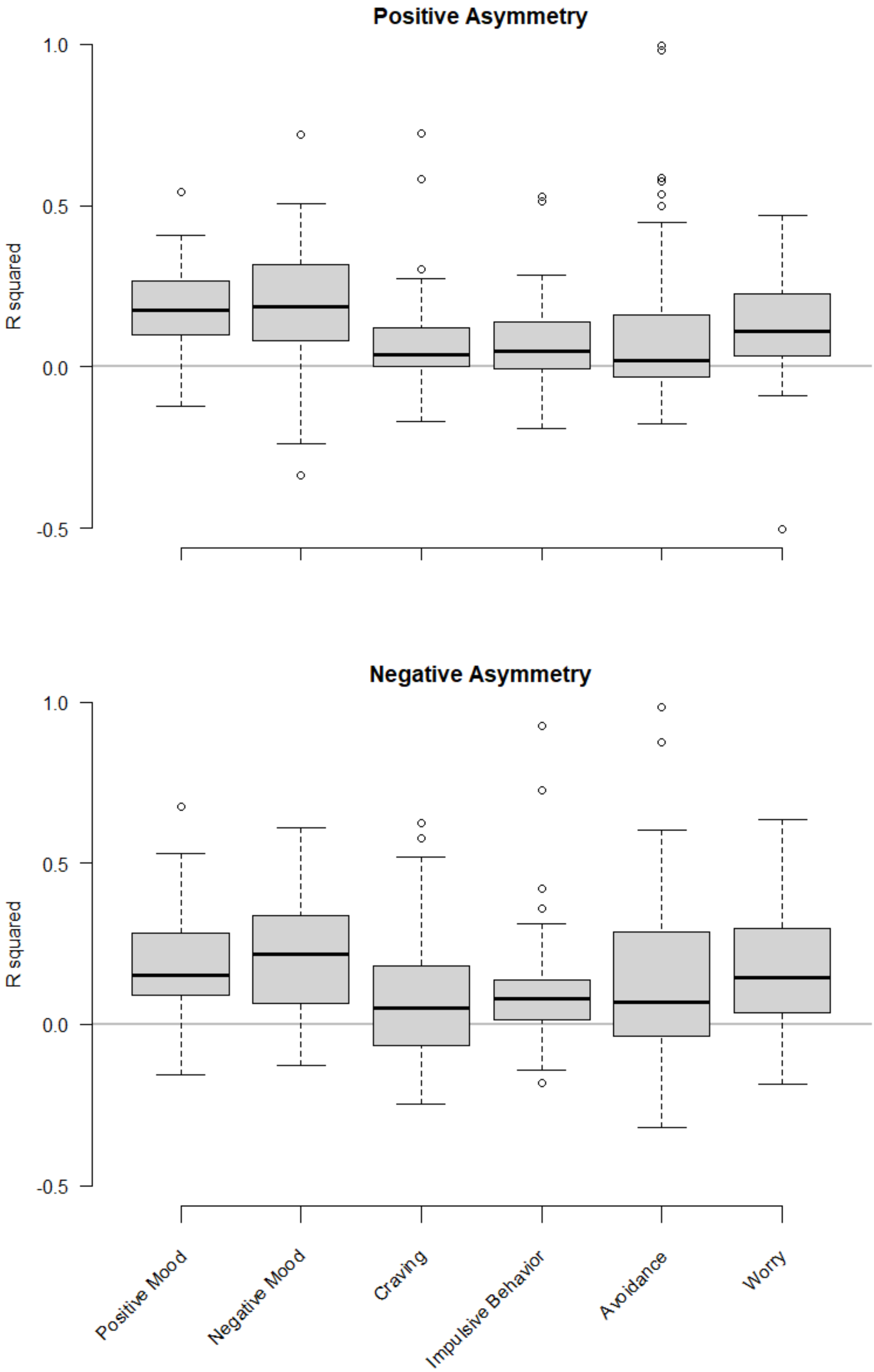
| Model | Variable | *B* | SE *B* | *β* | SE *β* | *t* | *p* | 95% CI for B | | Tolerance |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper | |
| Continuous | Positive Mood | -0.7 | 4.0 | 0.0 | 0.1 | -0.18 | .860 | -8.6 | 7.2 | .734 |
| | Negative Mood | -0.5 | 5.6 | 0.0 | 0.2 | -0.09 | .932 | -11.6 | 10.6 | .278 |
| | Craving | -9.7 | 7.1 | -0.2 | 0.1 | -1.37 | .174 | -23.7 | 4.3 | .614 |
| | Impulsive Behavior | 11.2 | 8.2 | 0.3 | 0.2 | 1.37 | .174 | -5.0 | 27.5 | .257 |
| | Avoidance | -5.3 | 6.8 | -0.1 | 0.2 | -0.78 | .440 | -18.8 | 8.3 | .318 |
| | Worry | -0.2 | 3.5 | 0.0 | 0.2 | -0.05 | .964 | -7.2 | 6.9 | .364 |
| Categorical | Positive Mood | 0.0 | 0.1 | 0.0 | 0.1 | -0.21 | .831 | -0.2 | 0.2 | .734 |
| | Negative Mood | 0.0 | 0.1 | 0.1 | 0.2 | 0.33 | .744 | -0.2 | 0.3 | .278 |
| | Craving | 0.2 | 0.2 | 0.2 | 0.1 | 1.35 | .181 | -0.1 | 0.5 | .614 |
| | Impulsive Behavior | -0.3 | 0.2 | -0.3 | 0.2 | -1.57 | .120 | -0.7 | 0.1 | .257 |
| | Avoidance | 0.1 | 0.1 | 0.1 | 0.2 | 0.62 | .537 | -0.2 | 0.4 | .318 |
| | Worry | 0.0 | 0.1 | 0.0 | 0.2 | 0.27 | .789 | -0.1 | 0.2 | .364 |

Note: CI: Confidence Interval.

*Model Validity*

Residual plots can be found in the supplementary files (https://osf.io/gwvxr/?view_only=5db25d85d7be4a7386e62566055b057e). When residuals deviated substantially from a normal distribution, the model typically included a large proportion of the same predictions in an attempt to fit a data distribution with very low variance. For example, a model fitted to a participant who reported no more than five instances of avoidance during the data collection period had predicted values barely deviating from the minimum, causing the residual graph to be dominated by a single peak. When a participant showed larger variance in their reporting, the residuals were closer to normally distributed. Low-variance responses were most common for craving, avoidance, and impulsive behavior.

The distribution of $R^2$ values between groups can be found in Supplementary Figure 1. Visual inspection suggests that the differences in model fit between groups are slight, with somewhat higher variance in negative mood in the positive asymmetry group and somewhat higher variance in worry in the negative asymmetry group. Notably, the avoidance node is characterized by high maximum $R^2$ values in both groups. Given that low variance responses to avoidance were common, this likely reflects models that accurately predict these low variance responses.




*Supplementary Figure 1: Boxplots of $R^2$ value distribution for each node, split by group.*

# Chapter 6

General Discussion

The goal of the current thesis was to investigate individual differences in the relative strength of appetitive and aversive learning, and how these differences are related to personality traits, symptoms of mental disorders, and networks of mental disorder symptoms. To this end, we developed a task in which we calculated a measure of learning asymmetry using both accuracy and confidence ratings and combined these in a difference score that indicated whether an individual is more sensitive to appetitive or aversive learning, known as learning asymmetry. This combined measure of sensitivity was hypothesized to play a role in multiple different mental disorders. In chapter 2, we analyzed the association between learning asymmetry and self-reports of neuroticism, impulsivity and anhedonia, to determine whether learning asymmetry measured individual differences that could be related to personality traits. In chapter 3, we modified the task for primary reinforcers and measured psychological distress and substance use in addition to impulsivity and anhedonia. In chapter 4, we returned to the original learning asymmetry task to further analyze the association with mental disorder symptoms and to test whether predictions following from reinforcement sensitivity theory were supported. Finally, in chapter 5, we investigated groups of participants with positive and negative learning asymmetry to test whether their symptom networks of mood and behavior differ significantly.

## Main Findings

In chapters 2, 3 and 4, the personality traits neuroticism, impulsivity, anhedonia, food reactivity, behavioral activation and behavioral inhibition[4] were measured. Those that were found to have a significant association with learning asymmetry were impulsivity, food reactivity, and behavioral inhibition, and those that did not were neuroticism, anhedonia, and behavioral activation, although participants high in neuroticism did show a general reduction in learning strength. For impulsivity, we found that more impulsive participants were less able to learn from aversive (but not appetitive) stimuli compared to less impulsive participants. Food reactivity showed a positive association with learning asymmetry, which, in contrast

[4] We refer to the BIS/BAS scale traits as such in this chapter to disambiguate them from UPPS-P impulsivity and DASS anxiety.

to impulsivity, was driven by participants high in food reactivity learning better from appetitive associations, but not aversive associations. Behavioral inhibition also showed a positive association with learning asymmetry, and participants scoring high on this trait had both improved appetitive learning and reduced aversive learning.

In chapters 3 and 4, associations between symptoms of psychopathology and learning asymmetry were studied. The variables studied were psychological distress (a composite measure of several symptom dimensions, see Derogatis & Melisaratos, 1983), substance use, depression, anxiety and stress. Significant associations were found between learning asymmetry and psychological distress as well as depression. We did not find significant associations for stress, anxiety, or substance use. Appetitive learning was shown to be weaker in participants with higher psychological distress, but aversive learning did not differ.

In chapter 5, network dynamics of mood, impulsive behavior, avoidance and craving were compared based on differences in learning asymmetry. Participants were divided into a positive and a negative asymmetry group, and the time-series effects within the networks of each group were compared. Results showed that in the positive asymmetry group, anger was followed by impulsive behavior significantly more frequently than in the negative asymmetry group. In addition, the autocorrelation for impulsive behavior was lower in the positive asymmetry group than in the negative asymmetry group, while the autocorrelation for avoidance was lower in the negative asymmetry group than in the positive asymmetry group. These autocorrelation differences suggest that, in addition to the effect of anger on impulsive behavior, the positive group showed a greater reactivity to impulsive behavior in general. Although the negative group did not show any inter-node effects that were not present in the positive group, the autocorrelation effect of avoidance indicates that they were overall more reactive to avoidance. More generally, the findings demonstrate that learning asymmetry may be one of the traits that influences sensitivity to mental disorders by moderating the dynamics of symptom networks.

## Learning Asymmetry and Individual Differences

While impulsivity was positively related to learning asymmetry in line with our hypothesis, it was unexpected that this relation was mainly due to differences in aversive learning. We expected that impulsivity would affect appetitive learning, given that impulsivity has been linked to drug dependence (Chamorro et al., 2012) and as such, impulsive individuals may be more sensitive to appetitive conditioning. Other studies have found that impulsivity is associated with attentional bias in appetitive conditioning studies (Hicks et al., 2015; Wardle et al., 2018) as well as desire to eat (van den Akker et al., 2013). Despite this, it appeared that impulsivity did not benefit the ability to acquire appetitive associations from the current results. However, there are other studies consistent with the finding that impulsivity is associated with weaker aversive learning: Patterson & Newman (1993) describe several studies showing that 'disinhibited'[5] participants were more error-prone in a Go/No-go task when both rewards and punishments were offered compared to when only punishments were offered, and in addition, they take less time to reflect on these errors. Therefore, it seems that the presence of appetitive outcomes makes impulsive participants less sensitive to aversive outcomes. This could explain why, in chapter 3, impulsivity was not associated with learning asymmetry: not only were appetitive and aversive USs each administered in separate blocks, but these USs were administered regardless of how they responded. In other words, the effect of impulsivity found in chapters 2 and 4 may be attributed to – under those specific experimental conditions – reduced sensitivity to aversive learning in more impulsive participants.

A similar explanation for this discrepancy was offered in chapter 3: we expected that the interaction between approach and avoidance tendencies (or the BAS and BIS) would be different in the case of intermixed appetitive and aversive trials compared to when they are presented in separate blocks. In chapter 4, we found that higher scores on the BIS were associated with a more positive learning asymmetry, to which both appetitive and aversive learning contributed, but not with the scores on any of the BAS subscales. However, if the BIS/BAS scale is intended to capture both trait impulsivity and trait anxiety, then neither of these results are in agreement with other such measures, such as the UPPS-P in chapters 2 and 4 and the Beck Anxiety Inventory as used by Shook et al. (2007), which was associated with weaker appetitive learning. This casts doubt on whether RST provides any explanation about this pattern of results and does not indicate that learning asymmetry is compatible with

[5] Their sample is characterized by extraversion, which they equate with being prone to disinhibition, or being impulsive.

the predictions made by RST. The BIS/BAS scale has recently received criticism regarding its continued use despite revised measures having shown more reliability (Espinoza Oyarce et al., 2022). Future research may therefore need to carefully consider the choice of self-report scales related to RST.

Meanwhile, the significant association between food reactivity and learning asymmetry, and specifically appetitive learning, suggests that there is, in fact, a trait associated with improved appetitive learning independent of impulsivity. This result also supports the notion that food sensitivity as measured by the PFS is part of a broader tendency towards appetitive learning including non-food stimuli (Lowe et al., 2009). That would imply that impulsivity as measured by the UPPS-P, which was shown to be unrelated to appetitive learning, is a distinct trait. However, no association between substance use and learning asymmetry was found, and thus we have no indication that a learning asymmetry in favor of appetitive learning is associated with maladaptive outcomes such as addiction. Similarly, there were no significant associations between learning asymmetry and neuroticism or anhedonia, which was not consistent with our hypotheses. The fact that anhedonia showed no association with learning asymmetry in either chapter 2 or chapter 3 is surprising seeing as it is one of the central symptoms of depression (APA, 2013). Given that several studies testing healthy volunteers showed an association between depression and weaker appetitive learning (Shook et al., 2007; Zbozinek et al., 2021), it may be that anhedonia was not sufficiently present in the current samples for an association with learning asymmetry to be detectable[6]. This indicates that other symptoms of depression may play a larger role in the reduction in appetitive learning. Conversely, neuroticism showed no association with learning asymmetry, only an overall reduction in learning: participants high in neuroticism showed both weaker appetitive and weaker aversive learning compared to those low in neuroticism. While this result is not strictly relevant to our research questions regarding learning asymmetry, it may be in line with Heinz et al. (2016) who suggested dysfunction in basic learning mechanisms may play a role in multiple different mental disorders. Further research regarding the role of neuroticism in learning may therefore be warranted.

[6] In the sample used in chapter 3, no participants had an average score lower than 2.9, on a scale of 1 to 4.

Overall, these results are encouraging, as they attest to learning asymmetry's potential to demonstrate relations with personality traits. As established in chapter 1, the rarity of an integrated approach including both appetitive and aversive learning in the field of conditioning research means that studies investigating individual differences may not properly contextualize their results when measuring only appetitive or aversive learning. However, there is one aspect of the integrative approach that has not proven as important as expected, namely that of calculating a difference score between appetitive and aversive learning. We anticipated both appetitive and aversive learning to be involved in producing the observed effects in learning asymmetry. However, this was only the case for inhibition, assessed with the BIS. For every other characteristic, either appetitive or aversive learning played only a very small role, and the significant association with learning asymmetry was entirely driven by the other component. In these cases, the use of learning asymmetry may not provide added value. For example, if aversive learning is associated with impulsivity, but not with appetitive learning, using learning asymmetry as the dependent variable in the analysis does not improve our ability to detect this association, and may in fact reduce it if learning asymmetry has higher variance than measures of appetitive or aversive learning. However, this does not imply that the approach of measuring both appetitive and aversive learning has the same drawback.

Assessing individual differences in sensitivity to appetitive versus aversive learning is comparatively understudied in conditioning research. Given the pattern of results that we have observed, the method of matching appetitive and aversive learning measures so that they can be subtracted from each other may not be most effective to determine the preferential association between a measure of individual differences and appetitive or aversive learning. If individual differences are most likely to be found in either appetitive or aversive learning, rather than both, it is less important that the two measures be as similar to each other as possible, as we have tried to achieve here. As this presented a major challenge for these studies, this is valuable knowledge for any future studies that plan to investigate appetitive and aversive learning. However, the strong versus weak situation (Lissek et al., 2006) should not be ignored: individual differences research continues to face other challenges. Namely, it is not only the methodology that can make the difference in detecting individual differences, but also in the choosing of the variables under study

(Hedge et al., 2018). For example, there may be substantial differences in inter-individual variation between skin conductance responses and expectancy ratings in response to an upcoming US. Even behavioral measures of appetitive and aversive learning have alternatives such as reaction time differences. Individual differences research therefore requires careful considerations of the expected variation in dependent variables.

**Learning Asymmetry and Mental Disorder Symptoms**

As discussed in chapter 1, the main aim of this thesis was to investigate whether deviations in basic mechanisms of learning are associated with symptoms of mental disorders. Dysfunctional learning is implicated in a wide variety of mental disorders and may therefore serve as a transdiagnostic mechanism, which may provide a better understanding of how mental disorders arise than attempting to classify each disorder according to which symptoms they are associated with (Heinz et al., 2016). In chapter 3, psychological distress was negatively associated with learning asymmetry. This is partially consistent with Shook et al. (2007), who found that weaker appetitive learning was associated with higher scores on self-report scales of depression and anxiety, and Zbozinek et al. (2021), who found that weaker appetitive learning was associated with higher trait depression. Examination of the BSI did not indicate that any subscales had a significantly stronger association with learning asymmetry than the others, suggesting that several different symptoms of psychopathology may have contributed to this effect. The same effect was not found in chapter 4, but these results were complicated by multicollinearity issues, making them less reliable. An alternative explanation for the presence of the effect of the BSI in chapter 3 and its absence in chapter 4 is that the use of primary reinforcers resulted in stronger learning effects as designed, but it is also possible that the classical versus instrumental conditioning procedures played a role. The task where participants were instructed to maximize their points may have been more motivating than the task where participants received the taste regardless of their response, which could have compensated for the expected reduction in appetitive learning in the participants who scored high on the BSI. The choice of US may also have been a factor, but differences in conditioned responses to primary and secondary reinforcers were previously shown to be comparable (Delgado et al., 2011). When investigating individual differences, however, variation in how participants respond to each may become more pronounced, leading to stronger learning effects. Regardless, taking into account previous studies, it appears that the most consistent result is that a negative learning asymmetry is associated with more symptoms of psychopathology, with depressive symptoms possibly playing the largest role in this effect.

However, we did not find any association between learning asymmetry and anxiety symptoms, which is unexpected given that reviews of anxiety disorder studies showed a relatively robust finding of impaired safety learning (Duits et al., 2015; Lissek et al., 2005). It is possible that the finding of weaker appetitive learning in higher psychological distress mirrors the weaker safety learning in anxiety disorder patients: safety learning can be considered a form of appetitive learning, as it stands opposite to the aversive stimuli used in these studies. As discussed in chapter 1, there is a stark contrast between the glut of aversive conditioning studies and the scarcity of appetitive conditioning studies on anxiety disorders. If anxiety disorder patients show weaker appetitive learning in general, in addition to weaker safety learning in aversive conditioning studies, then this suggests that appetitive learning ability may be very relevant to anxiety disorder in spite of being so rarely studied. Furthermore, it may serve as a transdiagnostic factor, being associated with anxiety, depression, and perhaps other mental disorders. Further study on clinical populations would be required to confirm this possibility.

Aside from psychological distress being associated with reduced appetitive learning, however, learning asymmetry did not show other associations with mental disorder symptoms. Food reactivity and impulsivity were both associated with learning asymmetry, and we expected that the positive asymmetry shown in both cases would be associated with increased susceptibility to substance use, but substance use was not found to be associated with learning asymmetry in any of our studies. In short, the significant associations between learning asymmetry and mental disorder symptoms appear to be one-sided. To echo the previous section, it thus seems unlikely that the approach of calculating a difference score between appetitive and aversive learning has added value over measuring either appetitive or aversive learning, in terms of its suitability as a transdiagnostic factor.

## Learning Asymmetry and Networks of Mood and Behavior Dynamics

In chapter 5, we compared the networks of a positive and a negative learning asymmetry group to determine whether they are differentially affected by network effects related to mood, impulsive behavior, avoidance and craving. Network analysis applied to psychological time series data is a relatively new method (Borsboom et al., 2021) that has been used in testing the network approach towards mental disorders (Fried et al., 2017). However, the feature space of variables describing mental states and behaviors that could be investigated in a network analysis is vast, making the selection of appropriate variables challenging. Despite this, measures of individual differences should not be overlooked. The network approach can be enriched by exploring the effects of individual differences such as learning asymmetry on symptom networks. Our network analyses showed that the mood and behavioral dynamics differed for the learning asymmetry groups: the positive asymmetry group shows greater reactivity to impulsive behavior while the negative asymmetry group shows greater reactivity to avoidance. This suggests that sensitivity to appetitive and aversive learning are traits that may be fruitful to investigate as part of the network approach.

The observed association between anger and impulsivity is one mechanism that may be further investigated for its relevance to mental disorders such as externalizing disorders. Such behavior has previously been observed in impulsive individuals (Cyders & Smith, 2008), and is included as the negative urgency subscale of the UPPS-P (Whiteside & Lynam, 2001), and as the UPPS-P has previously been shown to be associated with learning asymmetry, this indicates that the time series effects found in network analysis indeed correspond to actual differences in mood and behavior dynamics. However, this also raises the question of whether a network analysis is the most efficient way of investigating these effects. Compared to administering traditional self-report scales, EMA data collection is more intensive and time-consuming, and network analysis requires large sample sizes to achieve adequate statistical power. If existing self-report scales describe the same effects that are found using network analysis, it would be much more expedient to use only the self-report scales. The network hypothesis posits a role of individual differences in sensitivity to certain symptom interactions that can develop into a mental disorder (Borsboom, 2017), and the effect of anger on impulsive behavior may play a role in this. However, if network effects do not provide a more detailed explanation of behavior than existing self-report scales, the usefulness of network analysis for individual differences research may be limited.

**Methodological Considerations and Limitations**

In chapter 1, several modifications to the design of the learning asymmetry task were discussed, with the intent to improve earlier methods (Shook et al., 2007) of measuring individual differences in learning. Subsequent evaluations of the task have indicated that these were generally successful in maintaining the weak situation (Lissek et al., 2006) while at the same time enabling learning from multiple reinforcements. However, several limitations remain. Given that each task only measured learning asymmetry once per participant, it is unknown whether it is a stable trait or if it varies with multiple measurements. Determining the stability of learning asymmetry is especially challenging given that the task relies on complex figures that participants have not previously learned, and to obtain measures of learning asymmetry multiple times per participant would require an even greater variety of conditioned stimuli.

The task used in chapter 3 was notably different from the task used in the other empirical chapters. In the original version, participants made decisions about whether to accept objects that could result in gaining or losing points. Given that participants were able to affect whether they would gain or lose points, this resembled an instrumental conditioning procedure. This contrasts with the task using sweet and bitter tastes, which did not allow participants to affect whether they received the US, more akin to a classical conditioning procedure. In addition, the sweet and bitter tastes were each given in their own block, whereas the task using the point system intermixed gain and loss trials. Although the sweet and bitter taste version of the task was intended to elicit stronger learning from participants and thus allow for stronger effects of learning asymmetry to be measured, the differences between the two methods may ultimately have been too large to allow for comparable results. Given that psychological distress was only associated with learning asymmetry in chapter 3, and analyzing learning asymmetry as a combined measure did not show any advantages, future research into applying learning asymmetry as a transdiagnostic factor should proceed with using primary reinforcers as the US, although these do not necessarily need to be both taste stimuli.

A variety of self-report measures were used to determine the association of learning asymmetry with different traits relevant to personality and mental disorders. It is always challenging to select the appropriate measures to compare with learning asymmetry, as the options are myriad for both the traits to be examined and the self-report scales that examine them. The UPPS-P Impulsive Behavior Scale was chosen early on for the measurement of impulsivity, and as shown in chapter 2, was associated with learning asymmetry. Unfortunately, we were not aware until later that the UPPS-P is not designed to be averaged across all subscales, in contrast to, for example, the BSI. Consequently, the inclusion of the UPPS-P in subsequent experiments became necessary due to its association with learning asymmetry. However, this was done as a grand average, which deviated from the scale's intended purpose. Furthermore, considering that several other self-report scales were included in each analysis, analyzing all subscales separately would have diminished statistical power significantly. While this does not imply that the effect of impulsivity is unsound, it points to the possibility that alternative self-report measures would have been more appropriate to use.

Another complicating factor is the use of samples of healthy volunteers, rather than mental disorder patients. As learning asymmetry was investigated as a transdiagnostic factor, selecting, for example, Major Depressive Disorder (MDD) patients would not have been in line with the aims of the study. Instead, random variation in mental disorder symptoms was examined, with the assumption that if learning asymmetry was a transdiagnostic factor, a variety of symptoms would be associated with learning asymmetry. This proved not to be the case: only appetitive learning was weaker with high psychological distress, and while aversive learning was weaker with high impulsivity, this was not an effect that could be linked to mental disorder symptoms. Given that no other effects related to mental disorder symptoms were found, it is possible that other symptoms are associated with learning asymmetry, but they were insufficiently present in the samples tested here to detect.

Finally, it is noteworthy that the utilization of group comparisons in the network analysis presented in Chapter 5 was a compromise. By categorizing participants into positive and negative asymmetry groups, the variability in the learning asymmetry was reduced to a dichotomous variable, despite our explicit expectation that participants exhibiting a stronger learning asymmetry would also manifest more pronounced behavioral effects. Conversely, participants whose learning asymmetry was close to zero were assigned to the positive or negative group based on very small differences, while we expected them to be more alike in terms of behavior. However, the network analysis methods available at the time did not allow for the comparison of time series effects according to a continuous moderator variable, which necessitated the reduction of learning asymmetry to positive and negative groups. The analysis may therefore not have had sufficient statistical power to detect all relevant network effects.

**Future Research**

Given the characteristics of learning asymmetry thus presented, a general notion becomes clear: appetitive and aversive learning are both relevant traits in the field of individual differences research. Where appetitive learning is studied, aversive learning often goes neglected, and vice-versa. In a diverse range of domains, such as impulsivity and mental health, as well as unexplored areas, future research holds the potential to enhance our comprehension of how individual variations in learning influence behavior and mental well-being. This can be achieved by incorporating measures of appetitive and aversive learning into their research methodologies. This also has the potential to develop the theoretical basis of learning and behavior by elaborating on frameworks such as RST. Although the results from chapter 4 suggest that the predictions of RST do not align with the associations found with learning asymmetry, the inclusion of appetitive and aversive learning is crucial for a complete theory of learning and behavior. Thorough integrated learning studies are therefore highly recommended.

Aside from general recommendations regarding research on learning and behavior, the current results suggest a promising avenue of study in the appetitive learning differences found in both learning asymmetry and other studies on mental disorder symptoms. As discussed previously in this chapter, the association between weaker appetitive learning and psychological distress has parallels with the impaired safety learning found in anxiety disorder patients (Duits et al., 2015; Lissek et al., 2005). This is worth investigating not only for anxiety disorder patients, but also others such as MDD and possibly Obsessive-Compulsive Disorder (OCD) patients, as multiple symptom dimensions of the BSI may be involved. Of course, such a study should also measure aversive learning to ensure that the possibility of generally weaker learning, rather than weaker appetitive learning specifically, is not overlooked. If appetitive learning is indeed weaker in patients with different mental disorders, this would provide more evidence for impaired appetitive learning as a transdiagnostic factor. Several mechanisms are possible for this, such as excessive avoidance during safe situations or a reduced ability to recognize and approach rewarding opportunities. Regardless, such a finding would provide more opportunities for improving the treatment of mental disorders in ways that are not limited to rigid clinical diagnoses, and potentially enhance early detection and prevention of such disorders if appetitive learning is able to be targeted in interventions.

## Concluding Remarks

Although we did not conclusively demonstrate that measuring the relative strength of appetitive and aversive learning in a single variable has added benefits compared to using their component measures, the results of this thesis nevertheless show the added value of learning asymmetry. We have shown that learning asymmetry provides further insights into the interaction between appetitive and aversive learning and other individual differences that go overlooked when studies measure only appetitive or aversive learning. In addition, learning asymmetry can provide additional insight into the relation between learning and personality measures such as impulsivity, suggesting that these traits may have unexpected behavioral effects such as weaker aversive learning. Furthermore, the relation between learning asymmetry and psychological distress shows that learning asymmetry may be useful as a transdiagnostic factor, given that weaker appetitive learning is found in other studies on symptoms of mental disorders. Further research, both on individual differences and patients versus controls, is warranted to elucidate the comprehensive impact of appetitive and aversive learning disparities on the manifestation of mental disorder symptoms.

# Reference list

Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of Anxiety Disorders*, *44*, 36–46. https://doi.org/10.1016/j.janxdis.2016.10.003

Albrecht, U., Kirschner, N. E., & Grüsser, S. M. (2007). Diagnostic instruments for behavioural addiction: An overview. *GMS Psycho-Social Medicine*, *4*, Doc11.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (fifth ed.). American Psychiatric Publishing.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). American Psychiatric Publishing.

Andreatta, M., & Pauli, P. (2015a). Appetitive vs. Aversive conditioning in humans. *Frontiers in Behavioral Neuroscience*, *9*. https://www.frontiersin.org/articles/10.3389/fnbeh.2015.00128

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1990a). Flavor-flavor and color-flavor conditioning in humans. *Learning and Motivation*, *21*(4), 434–455. https://doi.org/10.1016/0023-9690(90)90025-J

Bakker, J. M., Goossens, L., Kumar, P., Lange, I. M. J., Michielse, S., Schruers, K., Bastiaansen, J. A., Lieverse, R., Marcelis, M., Van Amelsvoort, T., Van Os, J., Myin-Germeys, I., Pizzagalli, D. A., & Wichers, M. (2019). From laboratory to life: Associating brain reward processing with real-life motivated behaviour and symptoms of depression in non-help-seeking young adults. *Psychological Medicine*, *49*(14), 2441–2451. https://doi.org/10.1017/S0033291718003446

Banica, I., Schell, S. E., Racine, S. E., & Weinberg, A. (2022). Associations between different facets of anhedonia and neural response to monetary, social, and food reward in emerging adults. *Biological Psychology*, *172*, 108363. https://doi.org/10.1016/j.biopsycho.2022.108363

Beckers, T., Krypotos, A.-M., Boddez, Y., Effting, M., & Kindt, M. (2013b). What's wrong with fear conditioning? *Biological Psychology*, *92*(1), 90–96. https://doi.org/10.1016/j.biopsycho.2011.12.015

Borsboom, D. (2017a). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D. (2017b). Mental disorders, network models, and dynamical systems. In *Philosophical Issues in Psychiatry: Classification of psychiatric illness. IV* (pp. 80–97). Oxford University Press.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Briere, J., Hodges, M., & Godbout, N. (2010). Traumatic stress, affect dysregulation, and dysfunctional avoidance: A structural equation model. *Journal of Traumatic Stress*, *23*(6), 767–774. https://doi.org/10.1002/jts.20578

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), Article 2. https://doi.org/10.1037/0022-3514.67.2.319

Chamorro, J., Bernardi, S., Potenza, M. N., Grant, J. E., Marsh, R., Wang, S., & Blanco, C. (2012). Impulsivity in the general population: A national study. *Journal of Psychiatric Research*, *46*(8), 994–1001. https://doi.org/10.1016/j.jpsychires.2012.04.023

Charlson, F. J., Baxter, A. J., Dua, T., Degenhardt, L., Whiteford, H. A., & Vos, T. (2015). Excess mortality from mental, neurological and substance use disorders in the Global Burden of Disease Study 2010. *Epidemiology and Psychiatric Sciences*, *24*(2), 121–140. https://doi.org/10.1017/S2045796014000687

Clark, D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical Psychology*, *14*(Volume 14, 2018), 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Cooper, S. E., van Dis, E. A. M., Hagenaars, M. A., Krypotos, A.-M., Nemeroff, C. B., Lissek, S., Engelhard, I. M., & Dunsmoor, J. E. (2022). A meta-analysis of conditioned fear generalization in anxiety-related disorders. *Neuropsychopharmacology*, *47*(9), Article 9. https://doi.org/10.1038/s41386-022-01332-2

Coppin, G., Nolan-Poupart, S., Jones-Gotman, M., & Small, D. M. (2014a). Working memory and reward association learning impairments in obesity. *Neuropsychologia*, *65*, 146–155. https://doi.org/10.1016/j.neuropsychologia.2014.10.004

Corr, P. J. (2008). Reinforcement Sensitivity Theory (RST): Introduction. In *The reinforcement sensitivity theory of personality* (pp. 1–43). Cambridge University Press. https://doi.org/10.1017/CBO9780511819384.002

Cramer, A. O. J., Waldorp, L. J., Maas, H. L. J. van der, & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2–3), 137–150. https://doi.org/10.1017/S0140525X09991567

Cuijpers, P. (2019). Targets and outcomes of psychotherapies for mental disorders: An overview. *World Psychiatry*, *18*(3), 276–285. https://doi.org/10.1002/wps.20661

Cyders, M. A., & Smith, G. T. (2007). Mood-based rash action and its components: Positive and negative urgency. *Personality and Individual Differences*, *43*(4), 839–850. https://doi.org/10.1016/j.paid.2007.02.008

Cyders, M. A., & Smith, G. T. (2008). Emotion-based Dispositions to Rash Action: Positive and Negative Urgency. *Psychological Bulletin*, *134*(6), 807–828. https://doi.org/10.1037/a0013341

Daoud, J. I. (2017). Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series*, *949*, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Delgado, M., Jou, R., & Phelps, E. (2011). Neural Systems Underlying Aversive Conditioning in Humans with Primary and Secondary Reinforcers. *Frontiers in Neuroscience*, *5*. https://www.frontiersin.org/articles/10.3389/fnins.2011.00071

Delgado, M. R., Labouliere, C. D., & Phelps, E. A. (2006). Fear of losing money? Aversive conditioning with secondary reinforcers. *Social Cognitive and Affective Neuroscience*, *1*(3), 250–259. https://doi.org/10.1093/scan/nsl025

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, *13*(3), Article 3. https://doi.org/10.1017/S0033291700048017

Diaconescu, A. O., Jensen, J., Wang, H., Willeit, M., Menon, M., Kapur, S., & Mcintosh, R. (2011). Aberrant Effective Connectivity in Schizophrenia Patients during Appetitive Conditioning. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00239

Dingemans, A. E., Martijn, C., van Furth, E. F., & Jansen, A. T. M. (2009). Expectations, mood, and eating behavior in binge eating disorder. Beware of the bright side. *Appetite*, *53*(2), 166–173. https://doi.org/10.1016/j.appet.2009.06.002

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., van den Hout, M. A., & Baas, J. M. P. (2015a). Updated Meta-Analysis of Classical Fear Conditioning in the Anxiety Disorders. *Depression and Anxiety*, *32*(4), Article 4. https://doi.org/10.1002/da.22353

Elman, I., Lowen, S., Frederick, B. B., Chi, W., Becerra, L., & Pitman, R. K. (2009). Functional Neuroimaging of Reward Circuitry Responsivity to Monetary Gains and Losses in Posttraumatic Stress Disorder. *Biological Psychiatry*, *66*(12), 1083–1090. https://doi.org/10.1016/j.biopsych.2009.06.006

Espinoza Oyarce, D. A., Burns, R., Butterworth, P., & Cherbuin, N. (2022). A New Measure of the Revised Reinforcement Sensitivity Theory. *Social Psychology*, *53*(6), 327–340. https://doi.org/10.1027/1864-9335/a000502

Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral and Brain Sciences*, *2*(2), 188–199. https://doi.org/10.1017/S0140525X00061914

Eysenck, S. B. G., & Eysenck, H. J. (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder & Stoughton.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feil, M., Halvorson, M., Lengua, L., & King, K. M. (2020). A state model of negative urgency: Do momentary reports of emotional impulsivity reflect global self-report? *Journal of Research in Personality*, *86*, 103942. https://doi.org/10.1016/j.jrp.2020.103942

Field, M., & Cox, W. (2008). Attentional bias in addictive behaviors: A review of its development, causes, and consequences. *Drug and Alcohol Dependence*, *97*(1–2), 1–20. https://doi.org/10.1016/j.drugalcdep.2008.03.030

Finn, P. R., Kessler, D. N., & Hussong, A. M. (1994a). Risk for alcoholism and classical conditioning to signals for punishment: Evidence for a weak behavioral inhibition system. *Journal of Abnormal Psychology*, *103*(2), Article 2. https://doi.org/10.1037/0021-843X.103.2.293

Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 72. https://doi.org/10.1186/s12916-015-0325-4

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Gera, R., Barak, S., & Schonberg, T. (2019a). Counterconditioning following memory retrieval diminishes the reinstatement of appetitive memories in humans. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-45492-6

Gray, J. A. (1975a). *Elements of a Two-Process Theory of Learning*. Academic Press.

Gray, J. A. (1976). The behavioural inhibition system: A possible substrate for anxiety. In M. P. Feldman & A. M. Broadhurst (Eds.), *Theoretical and Experimental Bases of Behaviour Modification* (pp. 3–41). Wiley.

Haslbeck, J., Epskamp, S., & Waldorp, L. (2023). *Testing for Group Differences in Multilevel Vector Autoregressive Models*. OSF. https://doi.org/10.31234/osf.io/dhp8s

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Heinz, A., Schlagenhauf, F., Beck, A., & Wackerhagen, C. (2016). Dimensional psychiatry: Mental disorders as dysfunctions of basic learning mechanisms. *Journal of Neural Transmission*, *123*(8), 809–821. https://doi.org/10.1007/s00702-016-1561-2

Hicks, J. A., Fields, S., Davis, W. E., & Gable, P. A. (2015a). Heavy drinking, impulsivity and attentional narrowing following alcohol cue exposure. *Psychopharmacology*, *232*(15), 2773–2779. https://doi.org/10.1007/s00213-015-3912-3

Hicks, J. A., Fields, S., Davis, W. E., & Gable, P. A. (2015b). Heavy drinking, impulsivity and attentional narrowing following alcohol cue exposure. *Psychopharmacology*, *232*(15), Article 15. https://doi.org/10.1007/s00213-015-3912-3

Hoerger, M. (2013). *ZH: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations.* https://www.psychmike.com/dependent_correlations.php

Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., & Dolan, R. J. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine*, *46*(5), 1027–1035. https://doi.org/10.1017/S0033291715002597

Jansen, A., Broekmate, J., & Heymans, M. (1992). Cue-exposure vs self-control in the treatment of binge eating: A pilot study. *Behaviour Research and Therapy*, *30*(3), 235–241. https://doi.org/10.1016/0005-7967(92)90069-S

John, O. P., & Gosling, S. D. (2000). Personality traits. In *Encyclopedia of psychology, Vol. 6.* (2004-12704-044; pp. 140–144). Oxford University Press. https://doi.org/10.1037/10521-044

Jover Martínez, A., Lemmens, L. H. J. M., Fried, E. I., & Haslbeck, J. (2025). Does the structure of dynamic symptom networks depend on baseline psychopathology in students? *Doctoral Thesis*, 99–125.

Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., Sponheim, S. R., & Lissek, S. (2017). Neural Substrates of Overgeneralized Conditioned Fear in PTSD. *American Journal of Psychiatry*, *174*(2), 125–134. https://doi.org/10.1176/appi.ajp.2016.15121549

Kajanoja, J., & Valtonen, J. (2024). A Descriptive Diagnosis or a Causal Explanation? Accuracy of Depictions of Depression on Authoritative Health Organization Websites. *Psychopathology*, *57*(5), 389–398. https://doi.org/10.1159/000538458

Kangas, B. D., Der-Avakian, A., & Pizzagalli, D. A. (2022). Probabilistic Reinforcement Learning and Anhedonia. In D. A. Pizzagalli (Ed.), *Anhedonia: Preclinical, Translational, and Clinical Integration* (Vol. 58, pp. 355–377). Springer International Publishing. https://doi.org/10.1007/7854_2022_349

Kazak, A. E. (2018). Editorial: Journal Article Reporting Standards. *American Psychologist*, *74*(1), 1–2. https://doi.org/10.1037/amp0000263

Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. (2024a). Aversive conditioning is impaired in impulsive individuals: A study on learning asymmetries. *Journal of Behavior Therapy and Experimental Psychiatry*, *83*, 101939. https://doi.org/10.1016/j.jbtep.2023.101939

Kemp, L. T., Smeets, T., Jansen, A., & Houben, K. (2025). Distress is not delicious: Appetitive conditioning is weaker with high psychological distress. *Journal of Experimental Psychopathology*, *16*(1), 20438087251314526. https://doi.org/10.1177/20438087251314526

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011a). Counterconditioning: An effective technique for changing conditioned preferences. *Experimental Psychology*, *58*(1), 31–38. https://doi.org/10.1027/1618-3169/a000063

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., Üstün, T. B., & Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiology and Psychiatric Sciences*, *18*(1), 23–33. https://doi.org/10.1017/S1121189X00001421

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, *136*, 768–821. https://doi.org/10.1037/a0020327

Kruse, O., Tapia León, I., Stark, R., & Klucken, T. (2017). Neural correlates of appetitive extinction in humans. *Social Cognitive and Affective Neuroscience*, *12*(1), 106–115. https://doi.org/10.1093/scan/nsw157

Kwapil, T. R. (1998). Social anhedonia as a predictor of the development of schizophrenia-spectrum disorders. *Journal of Abnormal Psychology*, *107*, 558–565. https://doi.org/10.1037/0021-843X.107.4.558

Leganes-Fonteneau, M., Scott, R., & Duka, T. (2018). Attentional responses to stimuli associated with a reward can occur in the absence of knowledge of their predictive values. *Behavioural Brain Research*, *341*, 26–36. https://doi.org/10.1016/j.bbr.2017.12.015

Levitsky, D. A., & Shen, X. (2008). Food power scale predicts dessert eating, but not meal eating or portion size effect. *Appetite*, *51*(2), 381. https://doi.org/10.1016/j.appet.2008.04.147

Lissek, S., Pine, D. S., & Grillon, C. (2006b). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology*, *72*(3), 265–270. https://doi.org/10.1016/j.biopsycho.2005.11.004

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005b). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy*, *43*(11), 1391–1424. https://doi.org/10.1016/j.brat.2004.10.007

Lommen, M. J. J., Engelhard, I. M., & van den Hout, M. A. (2010). Neuroticism and avoidance of ambiguous stimuli: Better safe than sorry? *Personality and Individual Differences*, *49*(8), 1001–1006. https://doi.org/10.1016/j.paid.2010.08.012

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience & Biobehavioral Reviews*, *80*, 703–728. https://doi.org/10.1016/j.neubiorev.2017.07.007

Lowe, M. R., Butryn, M. L., Didie, E. R., Annunziato, R. A., Thomas, J. G., Crerand, C. E., Ochner, C. N., Coletta, M. C., Bellace, D., Wallaert, M., & Halford, J. (2009a). The Power of Food Scale. A new measure of the psychological influence of the food environment. *Appetite*, *53*(1), 114–118. https://doi.org/10.1016/j.appet.2009.05.016

Lowe, M. R., Butryn, M. L., Didie, E. R., Annunziato, R. A., Thomas, J. G., Crerand, C. E., Ochner, C. N., Coletta, M. C., Bellace, D., Wallaert, M., & Halford, J. (2009b). The Power of Food Scale. A new measure of the psychological influence of the food environment. *Appetite*, *53*(1), Article 1. https://doi.org/10.1016/j.appet.2009.05.016

Lynam, D. R., Smith, G. T., Whiteside, S. P., & Cyders, M. A. (2006a). The UPPS-P: Assessing five personality pathways to impulsive behavior. *Unpublished Report*.

Martin-Soelch, C., Linthicum, J., & Ernst, M. (2007). Appetitive conditioning: Neural bases and implications for psychopathology. *Neuroscience & Biobehavioral Reviews*, *31*(3), 426–440. https://doi.org/10.1016/j.neubiorev.2006.11.002

Masselink, M., van Roekel, E., Heininga, V. E., Vrijen, C., & Oldehinkel, A. (2019). *Domains Of Pleasure Scale (DOPS): Assessing pleasure across domains* [Preprint]. Open Science Framework. https://doi.org/10.31219/osf.io/bu7z5

McGrath, J. J., Lim, C. C. W., Plana-Ripoll, O., Holtz, Y., Agerbo, E., Momen, N. C., Mortensen, P. B., Pedersen, C. B., Abdulmalik, J., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Bromet, E. J., Bruffaerts, R., Bunting, B., De Almeida, J. M. C., De Girolamo, G., De Vries, Y. A., Florescu, S., … De Jonge, P. (2020). Comorbidity within mental disorders: A comprehensive analysis based on 145 990 survey respondents from 27 countries. *Epidemiology and Psychiatric Sciences*, *29*, e153. https://doi.org/10.1017/S2045796020000633

Meemken, M. T., Kube, J., Wickner, C., & Horstmann, A. (2018a). Keeping track of promised rewards: Obesity predicts enhanced flexibility when learning from observation. *Appetite*, *131*, 117–124. https://doi.org/10.1016/j.appet.2018.08.029

Meine, L. E., Müller-Bardorff, M., Recher, D., Paersch, C., Schulz, A., Spiller, T., Galatzer-Levy, I., Kowatsch, T., Fisher, A. J., & Kleim, B. (2024). Network analyses of ecological momentary emotion and avoidance assessments before and after cognitive behavioral therapy for anxiety disorders. *Journal of Anxiety Disorders*, *106*, 102914. https://doi.org/10.1016/j.janxdis.2024.102914

Meyer, M. D., Risbrough, V. B., Liang, J., & Boutelle, K. N. (2015a). Pavlovian conditioning to hedonic food cues in overweight and lean individuals. *Appetite*, *87*, 56–61. https://doi.org/10.1016/j.appet.2014.12.002

Moran, E. K., Culbreth, A. J., & Barch, D. M. (2022). Anhedonia in Schizophrenia. In D. A. Pizzagalli (Ed.), *Anhedonia: Preclinical, Translational, and Clinical Integration* (pp. 129–145). Springer International Publishing. https://doi.org/10.1007/7854_2022_321

Morriss, J., Biagi, N., Lonsdorf, T. B., & Andreatta, M. (2021). The role of intolerance of uncertainty in the acquisition and extinction of reward. *The European Journal of Neuroscience*, *53*(9), 3063–3071. https://doi.org/10.1111/ejn.15173

Ormel, J., Bastiaansen, A., Riese, H., Bos, E. H., Servaas, M., Ellenbogen, M., Rosmalen, J. G. M., & Aleman, A. (2013). The biological and psychological basis of neuroticism: Current status and future directions. *Neuroscience & Biobehavioral Reviews*, *37*(1), 59–72. https://doi.org/10.1016/j.neubiorev.2012.09.004

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., Rosmalen, J. G. M., & Oldehinkel, A. J. (2013). Neuroticism and Common Mental Disorders: Meaning and Utility of a Complex Relationship. *Clinical Psychology Review*, *33*(5), Article 5. https://doi.org/10.1016/j.cpr.2013.04.003

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J. D., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., … UnÜtzer, Jü. (2018). The *Lancet* Commission on global mental health and sustainable development. *The Lancet*, *392*(10157), 1553–1598. https://doi.org/10.1016/S0140-6736(18)31612-X

Patterson, C. M., & Newman, J. P. (1993a). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review*, *100*(4), 716–736. https://doi.org/10.1037/0033-295X.100.4.716

Patterson, C. M., & Newman, J. P. (1993b). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review*, *100*(4), Article 4. https://doi.org/10.1037/0033-295X.100.4.716

Piccirillo, M. L., & Rodebaugh, T. L. (2022). Personalized networks of social anxiety disorder and depression and implications for treatment. *Journal of Affective Disorders*, *298*, 262–276. https://doi.org/10.1016/j.jad.2021.10.034

Pickering, A. D., & Corr, P. J. (2008a). J.A. Gray's Reinforcement Sensitivity Theory (RST) of Personality. In G. Boyle, G. Matthews, & D. Saklofske, *The SAGE Handbook of Personality Theory and Assessment: Volume 1—Personality Theories and Models* (pp. 239–256). SAGE Publications Ltd. https://doi.org/10.4135/9781849200462.n11

Pietri, E. S., Fazio, R. H., & Shook, N. J. (2013). Weighting Positive Versus Negative: The Fundamental Nature of Valence Asymmetry. *Journal of Personality*, *81*(2), 196–208. https://doi.org/10.1111/j.1467-6494.2012.00800.x

Pitman, R. K., & Orr, S. P. (1986b). Test of the conditioning model of neurosis: Differential aversive conditioning of angry and neutral facial expressions in anxiety disorder patients. *Journal of Abnormal Psychology*, *95*(3), 208–213. https://doi.org/10.1037/0021-843X.95.3.208

Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. *Neuroscience & Biobehavioral Reviews*, *88*, 117–140. https://doi.org/10.1016/j.neubiorev.2018.03.015

Quarmley, M., Gur, R. C., Turetsky, B. I., Watters, A. J., Bilker, W. B., Elliott, M. A., Calkins, M. E., Kohler, C. G., Ruparel, K., Rupert, P., Gur, R. E., & Wolf, D. H. (2019). Reduced safety processing during aversive social conditioning in psychosis and clinical risk. *Neuropsychopharmacology*, *44*(13), 2247–2253. https://doi.org/10.1038/s41386-019-0421-9

Schechtman, E., Laufer, O., & Paz, R. (2010). Negative Valence Widens Generalization of Learning. *Journal of Neuroscience*, *30*(31), 10460–10464. https://doi.org/10.1523/JNEUROSCI.2377-10.2010

Sharpe, B. M., Simms, L. J., & Wright, A. G. C. (2021). Impulsivity, Affect, and Stress in Daily Life: Examining a Cascade Model of Urgency. *Journal of Personality Disorders*, *35*(4), 494–512. https://doi.org/10.1521/pedi_2020_34_465

Shook, N. J., Fazio, R. H., & Vasey, M. W. (2007b). Negativity bias in attitude learning: A possible indicator of vulnerability to emotional disorders? *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 144–155. https://doi.org/10.1016/j.jbtep.2006.10.005

Sjouwerman, R., Scharfenort, R., & Lonsdorf, T. B. (2020). Individual differences in fear acquisition: Multivariate analyses of different emotional negativity scales, physiological responding, subjective measures, and neural activation. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-72007-5

Smillie, L. D., Pickering, A. D., & Jackson, C. J. (2006b). The New Reinforcement Sensitivity Theory: Implications for Personality Measurement. *Personality and Social Psychology Review*, *10*(4), 320–335. https://doi.org/10.1207/s15327957pspr1004_3

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A Scale for the Assessment of Hedonic Tone the Snaith–Hamilton Pleasure Scale. *British Journal of Psychiatry*, *167*(1), 99–103. https://doi.org/10.1192/bjp.167.1.99

Sokol-Hessner, P., & Rutledge, R. B. (2019). The Psychological and Neural Basis of Loss Aversion. *Current Directions in Psychological Science*, *28*(1), 20–27. https://doi.org/10.1177/0963721418806510

Sperry, S. H., Lynam, D. R., & Kwapil, T. R. (2018). The convergence and divergence of impulsivity facets in daily life. *Journal of Personality*, *86*(5), 841–852. https://doi.org/10.1111/jopy.12359

Stansfeld, S., Clark, C., Bebbington, P. E., King, M., Jenkins, R., & Hinchliffe, S. (2016). Common Mental Disorders. In S. McManus, P. E. Bebbington, R. Jenkins, & T. Brugha (Eds.), *In: McManus, S and Bebbington, PE and Jenkins, R and Brugha, T, (eds.) Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. (Pp. 37-68). NHS Digital: Leeds, UK. (2016)* (No. 2; Issue 2, pp. 37–68). NHS Digital. http://content.digital.nhs.uk/catalogue/PUB21748/apms-2014-full-rpt.pdf

Tomko, R. L., Carpenter, R. W., Brown, W. C., Solhan, M. B., Jahng, S., Wood, P. K., & Trull, T. J. (2014). Measuring Impulsivity in Daily Life: The Momentary Impulsivity Scale. *Psychological Assessment*, *26*(2), 339–349. https://doi.org/10.1037/a0035083

van den Akker, K., Jansen, A., Frentz, F., & Havermans, R. C. (2013a). Impulsivity makes more susceptible to overeating after contextual appetitive conditioning. *Appetite*, *70*, 73–80. https://doi.org/10.1016/j.appet.2013.06.092

van den Akker, K., Jansen, A., Frentz, F., & Havermans, R. C. (2013b). Impulsivity makes more susceptible to overeating after contextual appetitive conditioning. *Appetite*, *70*, 73–80. https://doi.org/10.1016/j.appet.2013.06.092

van den Akker, K., Schyns, G., Breuer, S., van den Broek, M., & Jansen, A. (2019b). Acquisition and generalization of appetitive responding in obese and healthy weight females. *Behaviour Research and Therapy*, *123*, 103500. https://doi.org/10.1016/j.brat.2019.103500

van den Akker, K., Schyns, G., & Jansen, A. (2017b). Altered appetitive conditioning in overweight and obese women. *Behaviour Research and Therapy*, *99*, 78–88. https://doi.org/10.1016/j.brat.2017.09.006

van den Akker, K., Schyns, G., & Jansen, A. (2018). Learned Overeating: Applying Principles of Pavlovian Conditioning to Explain and Treat Overeating. *Current Addiction Reports*, *5*(2), Article 2. https://doi.org/10.1007/s40429-018-0207-x

van der Schaaf, M. E., Schmidt, K., Kaur, J., Gamer, M., Wiech, K., Forkmann, K., & Bingel, U. (2022b). Acquisition learning is stronger for aversive than appetitive events. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03234-x

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171–178. https://doi.org/10.1016/S2215-0366(15)00505-2

Vriends, N., Michael, T., Schindler, B., & Margraf, J. (2012). Associative learning in flying phobia. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(2), 838–843. https://doi.org/10.1016/j.jbtep.2011.11.003

Wardle, M. C., Lopez-Gamundi, P., & Flagel, S. B. (2018a). Measuring appetitive conditioned responses in humans. *Physiology & Behavior*, *188*, 140–150. https://doi.org/10.1016/j.physbeh.2018.02.004

Wardle, M. C., Lopez-Gamundi, P., & Flagel, S. B. (2018b). Measuring appetitive conditioned responses in humans. *Physiology & Behavior*, *188*, 140–150. https://doi.org/10.1016/j.physbeh.2018.02.004

Watson, D., Gámez, W., & Simms, L. (2005). Basic dimensions of temperament and their relation to anxiety and depression: A symptom-based perspective. *Journal of Research in Personality*, *39*, 46–66. https://doi.org/10.1016/j.jrp.2004.09.006

Watson, M. R., Voloh, B., Naghizadeh, M., & Womelsdorf, T. (2019b). Quaddles: A multidimensional 3-D object set with parametrically controlled and customizable features. *Behavior Research Methods*, *51*(6), 2522–2532. https://doi.org/10.3758/s13428-018-1097-5

Whiteside, S. P., & Lynam, D. R. (2001a). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*(4), 669–689. https://doi.org/10.1016/S0191-8869(00)00064-7

WHO ASSIST Working Group. (2002a). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, reliability and feasibility. *Addiction*, *97*(9), 1183–1194. https://doi.org/10.1046/j.1360-0443.2002.00185.x

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-17977-w

World Health Organization. (2011). *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level*. WHO, Geneva, Switzerland.

World Health Organization. (2022). *Mental disorders*. https://www.who.int/news-room/fact-sheets/detail/mental-disorders

Zbozinek, T. D., Wise, T., Perez, O. D., Qi, S., Fanselow, M. S., & Mobbs, D. (2021a). Pavlovian occasion setting in human fear and appetitive conditioning: Effects of trait anxiety and trait depression. *Behaviour Research and Therapy*, *147*, 103986. https://doi.org/10.1016/j.brat.2021.103986

Zhang, Z., Manson, K. F., Schiller, D., & Levy, I. (2014b). Impaired Associative Learning with Food Rewards in Obese Women. *Current Biology*, *24*(15), 1731–1736. https://doi.org/10.1016/j.cub.2014.05.075

# Appendix

## Publications and Presentations

*Peer-reviewed publications*

**Kemp, L. T.**, Smeets, T., Jansen, A., & Houben, K. (2024). Aversive conditioning is impaired in impulsive individuals: A study on learning asymmetries. *Journal of Behavior Therapy and Experimental Psychiatry*, *83*, 101939. https://doi.org/10.1016/j.jbtep.2023.101939

**Kemp, L. T.**, Smeets, T., Jansen, A., & Houben, K. (2025). Distress is not delicious: Appetitive conditioning is weaker with high psychological distress. *Journal of Experimental Psychopathology*, *16*(1), 20438087251314526. https://doi.org/10.1177/20438087251314526

*Submitted manuscripts*

**Kemp, L. T.**, Smeets, T., Jansen, A., & Houben, K. The balance of appetitive and aversive learning in personality and psychopathology. *Journal of Experimental Psychopathology.*

**Kemp, L. T.**, Smeets, T., Jansen, A., & Houben, K. Learning Asymmetry as a Predictor of Mood and Behavior Dynamics: A Network Analysis. *Journal of Behavior Therapy and Experimental Psychiatry.*

*Presentations at conferences and symposia*

**Kemp, L. T.** (2021) Learning asymmetries in the third dimension: measuring learning biases with Quaddles. Poster presented at the Annual EPP Day 2021, Egmond aan Zee, the Netherlands.

**Kemp, L. T.** (2021) Learning asymmetries in the third dimension: measuring learning biases with Quaddles. Poster presented at the NSMD Annual Meeting 2021, Haarlem, the Netherlands.

**Kemp, L. T.** (2023) Appetitive and aversive learning asymmetries: a disordered desires project. Oral presentation at the EPP PhD Day, Utrecht, the Netherlands.

**Kemp, L. T.** (2024) Learning asymmetry group network effects: Do appetitive and aversive learning differences affect EMA responses? Poster presented at the NSMD Annual Meeting 2024, Maastricht, the Netherlands.

**Kemp, L. T.** (2025) The Good, the Bad, and the Networks: appetitive and aversive learning in the balance of mental health. Oral presentation at the NSMD Annual Meeting 2025, Maastricht, the Netherlands.

**Summary**

This thesis explores several approaches towards the study of (maladaptive) behavior and symptoms of mental disorders related to appetitive and aversive learning. These approaches are centered on a relative measure of sensitivity to appetitive versus aversive learning, known as the learning asymmetry. This measure is defined as positive when someone learns better from appetitive than aversive reinforcement, and negative when someone learns better from aversive than appetitive reinforcement. Learning asymmetry is investigated through correlational designs for its association with personality traits and symptoms of psychopathology, and through network models comparing positive and negative asymmetry groups on networks including mood states, impulsive behavior, craving, and avoidance. Using these methods, we aim to determine whether learning asymmetry provides better predictive value for mental disorder symptoms than comparable approaches that are limited to either appetitive or aversive learning, and whether differences in learning asymmetry are predictive of specific mood and behavioral dynamics as determined by network analysis.

The network hypothesis suggests that mental disorders can be best understood as symptom networks, which can reach a disordered state when multiple symptoms mutually reinforce each other, and the set of these symptoms that end up in a disordered state can be described as a mental disorder. Consequently, the network hypothesis argues that to effectively study and treat mental disorders, we must study how different symptoms interact with each other over time. It has also been suggested that individual differences in network connectivity play a role in how vulnerable individuals are to certain mental disorders, which indicates a possible role of traits such as learning asymmetry in the development and maintenance of symptoms of mental disorders.

Through investigating self-report measures of personality traits, it was shown that a more positive learning asymmetry was associated with higher impulsivity (Chapters 2, 4) and food reactivity (Chapter 4). While the association with impulsivity was driven primarily by participants with higher impulsivity showing reduced aversive learning, the reverse was true for food reactivity, showing that participants with higher food reactivity had improved appetitive learning. This shows an

interesting contrast between traits commonly associated with sensitivity to reward, and demonstrates that learning asymmetry is able to reveal contrasting associations with different personality traits. Furthermore, a more negative learning asymmetry was associated with higher psychological distress, a composite measure of several symptom dimensions that occur in common mental disorders (Chapter 3). This effect was driven by reduced appetitive learning in participants who scored high on psychological distress. This is noteworthy as several previous studies have demonstrated a similar effect, including reviews on anxiety disorder showing a reliable effect of reduced safety learning. This suggests that reduced appetitive (relative to aversive) learning may play a role in several different disorders, and may indicate a potential transdiagnostic mechanism by which these disorders develop. Further research investigating both appetitive and aversive learning in the context of mental disorders is therefore warranted.

The hypothesis that differences in learning asymmetry are associated with differences in mood and behavior dynamics was tested by comparing groups of positive and negative asymmetry individuals by their networks, which visualized the interactions between positive mood, negative mood, worry, impulsive behavior, craving, and avoidance over time in their day-to-day lives (Chapter 5). In the positive asymmetry group, self-reports of anger predicted impulsive behavior, whereas this effect was not present in the negative asymmetry group. Furthermore, the positive asymmetry group was overall more reactive to impulsive behavior, while the negative asymmetry group was overall more reactive to avoidance. Notably, the effect of anger predicting impulsive behavior is known to be associated with trait impulsivity, and is described as negative urgency, or a tendency to act rashly in response to negative mood. These results show that differences in learning asymmetry are indeed associated with differences in network effects, suggesting that network models can be used in conjunction with individual differences to better understand their effect on symptom networks. As the network hypothesis suggests, such differences in symptom networks may provide clues towards vulnerability to mental disorders.

The conclusions of this thesis emphasize the benefits of an integrative approach towards the study of appetitive and aversive learning and its relation to mental disorder symptoms. Although the specific implementation of learning asymmetry in the studies described here has some drawbacks, this does not negate that our understanding of the relation between appetitive and aversive learning on the one hand, and psychopathology on the other, can be much improved by examining individual differences in sensitivity to learning. Likewise, the network approach can be enriched by investigating how symptom networks vary according to individual differences, of which learning asymmetry is one example.

## Samenvatting

Dit proefschrift gaat over de studie van (maladaptief) gedrag en symptomen van mentale stoornissen, en hoe deze samenhangen met individuele verschillen in appetitief en aversief leren. Het proefschrift omvat verschillende benaderingen van de vraagstukken rondom dit onderwerp, die draaien om een bepaalde maat van individuele gevoeligheid voor appetitief en aversief leren, namelijk de leerasymmetrie. Deze wordt gedefinieerd als positief wanneer iemand beter appetitief dan aversief leert, en negatief wanneer iemand beter aversief dan appetitief leert. Eerst wordt hier de correlatie tussen leerasymmetrie enerzijds en persoonlijkheidskenmerken en symptomen van mentale stoornissen anderzijds onderzocht. Vervolgens worden netwerkmodellen gebruikt om groepen met een positieve dan wel negatieve leerasymmetrie met elkaar te vergelijken via de netwerk-effecten van gemoedstoestand, impulsief gedrag, verlangen, en vermijding. Hiermee wordt bepaald of leerasymmetrie een betere voorspeller is van symptomen van mentale stoornissen dan vergelijkbare maten die alleen appetitief of aversief leren in acht nemen, en of verschillen in leerasymmetrie indicatief zijn van de aanwezigheid van bepaalde interacties tussen gemoedstoestand en gedrag, zoals aangegeven wordt door de netwerkanalyse.

Volgens de netwerkhypothese kunnen mentale stoornissen het beste verklaard worden als netwerken van symptomen, die in een ontregelde staat kunnen komen wanneer meerdere symptomen elkaar versterken. De netwerkhypothese stelt vervolgens dat een mentale stoornis gedefinieerd wordt als een verzameling van elkaar versterkende symptomen die samen een ontregeld netwerk vormen. Om zulke mentale stoornissen te begrijpen en behandelen, stelt deze hypothese dat de interacties tussen symptomen gedurende langere tijd onderzocht moeten worden door middel van een netwerk. Ook is voorgesteld dat individuele verschillen in de connectiviteit van het netwerk van symptomen een rol spelen in de kwetsbaarheid van mensen voor bepaalde mentale stoornissen. Kenmerken zoals leerasymmetrie kunnen daarom relevant zijn voor de ontwikkeling en het behoud van mentale stoornissen.

Met vragenlijsten over persoonlijkheidskenmerken werd aangetoond dat een meer positieve leerasymmetrie verband had met hogere impulsiviteit (Hoofdstukken 2, 4) en reactiviteit voor eten (Hoofdstuk 4). Enerzijds bleek dat het effect van impulsiviteit gedreven werd door een afname in aversief leren in proefpersonen die hogere impulsiviteit rapporteerden; impulsiviteit bleek dus gerelateerd aan slechter aversief leren. Anderzijds was het effect van reactiviteit voor eten gedreven door een toename in appetitief leren in proefpersonen met hogere reactiviteit voor eten; reactiviteit voor eten bleek dus gerelateerd aan beter appetitief leren. Hiermee wordt een interessant contrast aangetoond, aangezien het hier gaat om twee kenmerken (impulsiviteit en reactiviteit op eten) die beide geassocieerd zijn met gevoeligheid voor beloningen. Dit leert ons dat het mogelijk is om contrasterende associaties met verschillende persoonlijkheidskenmerken te onderscheiden door gebruik van leerasymmetrie. Verder had leerasymmetrie een negatief verband met psychische nood, een gecombineerde maat van symptomen die voorkomen in algemene mentale stoornissen (Hoofdstuk 3). Dit effect werd gedreven door een afname in appetitief leren in proefpersonen met hoge psychische nood. Dit is opvallend aangezien eerder onderzoek een vergelijkbaar effect hebben aangetoond: recensies van studies naar patiënten met angststoornissen laten een robuust effect zien van verminderd leren van veiligheidssignalen. Dit geeft aan dat verminderd appetitief leren (relatief tot aversief leren) een rol kan spelen in meerdere mentale stoornissen, en er mogelijk een transdiagnostisch mechanisme bestaat dat invloed heeft op hoe deze stoornissen zich ontwikkelen. Er is daarom goede reden voor meer onderzoek naar de interactie tussen appetitief en aversief leren in de context van mentale stoornissen.

De hypothese dat verschillen in leerasymmetrie in verband staan met interacties tussen gemoedstoestand en gedrag werd getest door groepen van proefpersonen met positieve dan wel negatieve leerasymmetrie met elkaar te vergelijken. Dit werd gedaan door middel van hun netwerken, die de interacties tussen positieve emoties, negatieve emoties, zorgen, impulsief gedrag, verlangen, en vermijding visualiseren in het verloop van het dagelijkse leven (Hoofdstuk 5). In de groep met positieve asymmetrie voorspelde woede impulsief gedrag, wat niet het geval was in de groep met negatieve asymmetrie. Verder was de groep met positieve asymmetrie in het algemeen meer reactief op het gebied van impulsief gedrag. Daarentegen was de groep met negatieve asymmetrie meer reactief op het gebied van vermijding. Opmerkelijk was dat het effect van woede op impulsief gedrag een bekend fenomeen is in de studie van het kenmerk impulsiviteit, waar het beschreven wordt als negatieve urgentie, of een neiging om zich roekeloos te gedragen als gevolg op negatieve emoties. Deze resultaten tonen aan dat verschillen in leerasymmetrie inderdaad verband houden met verschillen in netwerkeffecten, wat aangeeft dat deze netwerkmodellen gebruikt kunnen worden in combinatie met individuele verschillen om hun effect op symptoomnetwerken beter te begrijpen. Zoals de netwerkhypothese voorstelt kunnen zulke verschillen in symptoomnetwerken aanwijzingen geven over hoe mensen kwetsbaar zijn voor mentale stoornissen.

De conclusies van dit proefschrift benadrukken de baten van een geïntegreerde aanpak bij de studie van appetitief en aversief leren, en hun verband met symptomen van mentale stoornissen. Hoewel er bepaalde nadelen zijn aan de implementatie van leerasymmetrie in de studies beschreven in dit proefschrift, neemt dit niet weg dat het onderzoeken van individuele verschillen in affiniteit voor leren belangrijk is voor ons begrip van de relatie tussen enerzijds appetitief en aversief leren, en anderzijds mentale stoornissen. Eveneens kunnen methoden met netwerkmodellen verrijkt worden door symptoomnetwerken te onderzoeken in de context van individuele verschillen, waarvan leerasymmetrie één voorbeeld is.

## Impact Addendum

### Relevance

Among all diseases, mental disorders are one of the greatest problems burdening society (Vigo et al., 2016), causing not only long-term suffering and distress but also many premature deaths (Charlson et al., 2015). A large portion of the global population suffers from a mental disorder in their lifetime, between 18 and 36% worldwide (Kessler et al., 2009). Many of these people fail to receive treatment: 35% of mental disorder patients worldwide are not treated (WHO, 2011). Attempts to treat mental disorders are often underwhelming: treatments often fail or lose their effectiveness over time, causing patients to relapse (Clark, 2018). The limited success of effectively treating mental disorders suggests that they are poorly understood (Cuijpers, 2019).

The standard approach to diagnose and treat mental disorders is based on the Diagnostic and Statistical Manual of Mental Disorders (DSM), which views mental disorders as separate things, an approach which is widely criticized (Cuijpers, 2019). Not only do patients who share the same disorder vary widely in their symptoms (Fried & Nesse, 2015), a mental disorder patient is often diagnosed with more than one disorder, which is a result of how the symptoms of many disorders are related and overlapping (Cramer et al., 2010). This creates many difficulties for the treatment of mental disorders from the approach of the DSM.

An alternative method of studying mental disorders is known as the network approach. This approach seeks to understand mental disorders as the result of interactions between symptoms, specifically in the form of a network where positive and negative relations between symptoms are visualized, in order to better understand how these symptoms may contribute to a disordered mental state. In this thesis, we demonstrated that network analysis can reveal differences in network structures between individuals who have different sensitivity to reward and punishment. These findings indicate that network analysis can be used to study the relation between individual differences and sensitivity to mental disorder symptoms, as well as how they influence network effects of mood and behavior. Furthermore, these findings show that several personality traits and mental disorder symptoms may be associated with sensitivity to reward and punishment. For this reason, the interaction between reward and punishment sensitivity warrants further study, as it may contribute to our understanding of maladaptive behaviors that impact mental health. These findings are relevant for clinical psychologists, people struggling with mental health problems, and policymakers.

**Target groups**

The first target group for this thesis is clinical psychologists. The network approach to mental disorders is growing in popularity, but the role of personality differences in network models is relatively understudied. Clinical psychologists interested in the significance of personality differences to mental disorders may be interested in applying network analysis to create a better understanding of how mental disorders and personality traits may play a role in sensitivity or resilience to mental disorders. With greater insights into symptom networks and personality differences, future studies may make further strides towards effective interventions for people suffering from mental disorders or prevention targeted at people who have a higher risk of developing a mental disorder.

The second target group is people struggling with mental disorders. Research has shown that using a mental disorder diagnosis, such as depression, to explain the symptoms a person might experience can create misconceptions among laypeople, who may be convinced that their symptoms are caused by a factor they have no control over (Kajanoja & Valtonen, 2024). Such beliefs can have negative consequences like maladaptive behavior regarding their symptoms and a loss of agency towards their own recovery. The approach of representing a mental disorder as a network of interacting symptoms, as in chapter 5, may provide people with mental disorders with a better understanding of how their day-to-day behavior causes their mental health to improve or deteriorate. In addition, the approach of studying individual differences in sensitivity to mental disorder symptoms may allow people with mental disorders to consider their individual nature and how this may affect any symptoms they might be living with.

The third target group is policymakers. The theories and models presented in this thesis may be relevant to policies regarding clinical practice and research, and the results suggest several areas of research that may be prioritized to achieve more effective treatment for mental disorders. Specifically, the differences between networks of mood and behavioral dynamics shown in this thesis suggest that network analysis has significant potential to reveal interactions between symptoms and behavior that may enable more effective interventions for mental disorders, and that individual differences such as sensitivity to reward and punishment are potentially valuable for exploring how individuals may be vulnerable to different types of mental disorder. Policymakers may therefore want to prioritize research that applies a transdiagnostic approach to mental disorder research, and which explores potential mechanisms by which mental health turns to disorder.

**Activities**

At the time of writing, chapter 2 has been published in Journal of Behavior Therapy and Experimental Psychiatry, and chapter 3 in Journal of Experimental Psychopathology. Chapter 4 has been submitted to Journal of Experimental Psychopathology, and chapter 5 to Journal of Behavior Therapy and Experimental Psychiatry. Results were also presented at the annual meeting of the Experimental Psychopathology postgraduate school, and the annual meeting of the New Science of Mental Disorders consortium, both of which are organizations of researchers from across the Netherlands and Belgium.

## Curriculum Vitae

Laurens Tristan Kemp was born in Zuid Scharwoude, the Netherlands, on 27 November 1990. He graduated in 2009 from the Christelijke Scholengemeenschap Jan Arentsz in Alkmaar. Afterwards, he started a bachelor's in biology at the University of Amsterdam, but switched to psychobiology in that same year. He obtained a bachelor's degree in psychobiology in 2012, and continued to enroll in the Research Master Brain and Cognitive Sciences (Cognitive Neuroscience track) at the University of Amsterdam, during which he completed internships at the Developmental Psychology department of the University of Amsterdam and the Centre for Research on Consciousness and Anomalous Psychology at Lund University in Sweden. Having graduated in 2015, he spent some time taking care of pets in Amsterdam, and later accepted a PhD position in 2020 at the University of Maastricht, supervised by Dr. Katrijn Houben, Prof. Dr. Tom Smeets, and Prof. Dr. Anita Jansen. In the future he plans to continue working for the benefit of Open Science, building on his work as part of the Free Our Knowledge project.

## Acknowledgements

First I would like to thank my supervisors: Katrijn, for being my reliable first point of contact and always being ready with advice, encouragement and support; Anita, for your trust and understanding, for being the driving force behind this project, and for ensuring that I was well-supported as a PhD student along with the others at CPS; and Tom, for your expertise and valuable feedback. My thanks also to the committee members for their time and effort spent in evaluating this thesis.

My heartfelt thanks goes out to the people at the UM who made going to the office a joy in the past three-and-a-bit years: my various officemates over the years, namely Leo, Michelle, Haixu, and both Anouks; those who were always great company over coffee or lunch, namely Alberto, Joey, Liyang, Hanna, Li, Mila, Esmée, Maaike, Anda, Codrin, Gudrun, Gwen, and Ophelia; and the Eatgroup, who set a standard with their combined professionalism and camaraderie that I can only hope my future colleagues live up to.

Thank you to the NSMD members who have been a source of support over the years: Jonas and Lourens for their tireless efforts in educating us and helping us with the network analysis methods, which allowed the final paper of this thesis to take shape; likewise the PhD students who all made an effort to have online meetings and discussions in spite of being spread all over the country, who made it feel less like I was thrown into the deep end of the network analysis pool: Gita, Myrthe, Inga Marie, Mado, Iara, and Franzi; thanks also to everyone else who made the NSMD events and activities productive and fun.

Thank you to my good friends Koen, Dylan, Ellen, Minha, Tessa, Huub, Tim, Serena and Scott. You are still my favorite people and I'm very grateful for our continued friendship. Thank you to Emma, my sister, for always being there to talk to. I am grateful to my parents, Frits and Joke, for their work in getting me to where I am. I know they would be proud. Last but not least, thank you to my lovely partner, Louella, for being the greatest source of peace in my life.