

Research Articles

Measures of Implicit Cognition for Marketing Research

By Thorsten Teichert, Alexander Graf, Sajad Rezaei, Philipp Wörfel and Helen Duh

Automatic, unconscious processes largely influence human decision-making. However, quantitative market research focuses on eliciting conscious responses. This foregoes the opportunity to investigate – and steer – preceding cognitive processes of decision-making. Three implicit cognitions are of special relevance along consumers' journey: Implicit attention introduces the first perception of a stimulus. Implicit associations can cause attitude and preference formation. Finally, approach tendencies can induce impulse buying. This paper provides a broad methodological overview of these implicit cognition measures to guide future researchers' marketing applications. It presents the methods' theoretical foundations, outlines how they can overcome explicit measures' limitations, and sketches their potential for marketing applications. In addition, the authors describe important research paradigms, alternative experimental setups, and data analyses steps to enable researchers to use implicit measurement tools. The measurement instruments are implemented in a non-profit software (AskYourBrain) and tested in an illustrative study. Key findings are summarized and marketing application suggestions made.

1. Introduction

Decades of social psychology research show that most decision-making and social behaviour occur without the individual being aware of them (Hofmann et al. 2005). Even economists accept the influence of automatic processes that the brain's intuitive system drives (Kahneman 2003). Despite this acceptance, a cognitive approach, according to which decisions are intentionally made, still largely dominates marketing research and, specifically, branding and new product design research (Bargh 2002). This paper aims at bridging this gap in market research by examining three complementary measurement instruments that measure implicit cognition's key facets quantitatively.

A severe drawback of implicit cognition measures had long been that they required a researcher's physical presence, or that their usage was restricted to laboratory settings. Currently, online implementations of implicit cognition measurements enable a cost-efficient and timely execution of large-scale and representative surveys in the implicit cognition research field. Playful cognition experiments can be executed in mobile devices and allow for real-time feedback. This addresses another common cognition measurement problem, as the tasks are usually viewed as repetitive, boring, and frustrating. Respondents find implicit cognition measures less burdensome and convey a "touch and feel" experience, thereby pro-



Thorsten Teichert holds the Chair of Marketing and Innovation, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, Phone: +49/40 42838 4643, Fax: +49/40 42838 5250, E-mail: thorsten.teichert@uni-hamburg.de
*Corresponding author



Alexander Graf is research assistant and doctoral student at the Chair of Marketing and Innovation, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, Phone: +49/40 42838 7836, Fax: +49/40 42838 5250, E-mail: alexander.graf@uni-hamburg.de.



Sajad Rezaei is postdoctoral researcher at the Chair of Marketing and Innovation, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, Phone: +49/40 42838 9095, Fax: +49/40 42838 5250, E-mail: sajad.rezaei@uni-hamburg.de.



Philipp Wörfel is research assistant and doctoral student at the Chair of Marketing and Innovation, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, Phone: +49/40 42838 7835, Fax: +49/40 42838 5250, E-mail: philipp.woerfel@uni-hamburg.de.



Helen Duh is Senior Lecturer at the Department of Marketing, School of Economic & Business Sciences, University of the Witwatersrand Johannesburg, 1 Jan Smuts Avenue, Johannesburg, 2000, South Africa, Phone: +27/117178217, Fax: +27 11 716 5202 E-mail: helen.du@wits.ac.za

viding the opportunity to reach out to a broader audience, especially to younger generations (Lumsden et al. 2016). In order to provide insights into implicit cognition research's new and promising paradigm, we specifically report on a non-profit software implementation in the illustrative application.

The objectives of this study are fourfold: (a) to provide a broad overview of important implicit cognition measures, (b) to initiate ideas for future marketing applications, (c) to provide guidance for designing and executing such studies, and (d) to critically consider the shortcomings and the methodological issues. We therefore focus on three important methods that we consider especially relevant for marketing applications.

1.1. Theoretical Background

Human cognition encompasses a broad range of mental processes, such as perception, memory, mental imagery, problem solving, and carrying out actions (Brandimonte et al. 2006). Simplified models of human cognition differentiate between two separable, but interrelated, systems in which cognitions are processed differently (see overview in: Gawronski and Bannon 2019). A series of well-recognized experiments (for a review, see Camerer et al. 2004) that behavioural economists conducted, demonstrate that purely rational and highly reflective decision-making processes – which the Nobel prize laureate, Kahneman (2003) labels “system 2” processes – account for only a minor share of human decision-making. Quasi-automatic, less conscious, and hardly cognitively reflected processes – labeled “system 1” by Kahneman – steer daily decision-making. Such violations of the “homo economicus” are especially prevalent with regard to typical consumption decisions, whether regarding impulsive shopping, hedonic goods or other fast-moving consumer goods, such as (unhealthy) foods or beverages.

Dual process models have far-reaching roots in psychology (Gawronski and Bannon 2019). A number of decades ago, Graf and Schacter (1985) introduced the well-known implicit/explicit dichotomy in their research on long-term memory. These scholars defined implicit memory and explicit memory as distinct retrieval processes of information stored in the memory, the former being unconscious and not deliberate, the latter being conscious and intentional. Social psychology has revealed that traces of past experience can affect behaviour even if this earlier experience is not available for self-reporting or introspection (Greenwald and Banaji 1995). Consequently, past learning episodes and consumption experiences affect evaluations and behaviour without the individual being aware of their influential effect. In contrast to the discourse on “hidden persuaders”, the constituent proposition is unawareness of the cognitive process and not of the stimuli as such (Stacy and Wiers 2010). Implicit cognition processes can also interact with explicit processes in various ways (e. g., Sun et al. 2005), even in compulsive consumption situations (Thush et al. 2008).

An implicit cognitive process does not require any effort from the individual and is initiated automatically without the individual being able to avoid a reaction when exposed to the appropriate stimuli. There are various implicit processes and, accordingly, a multitude of different operationalisations (De Houwer et al. 2009; De Houwer and Moors 2010). Recent research on neuroscientific cognition differentiates between three types of implicit cognitive processes (Stacy and Wiers 2010) with special relevance for consumer decision-making:

- An attentional bias, whereby cognitions are activated without conscious control, characterize implicit attention. This leads to a strengthened perception of the underlying stimulus (Waters and Feyerabend 2000). For example, a confrontation with a fast food chain logo can automatically stimulate hungry fast-food lovers' attention, whereas the brand's non-likers might not even perceive the displayed stimulus.
- Implicit associations build a second layer of implicit cognitions, which a stimulus presentation can automatically evoke (Anderson 1983). For example, mentioning the word “Apple” can provoke spontaneous positive associations with “fun” or “innovative” in individuals familiar with the brand. Such associations and the resulting emotions are, in general, not consciously reflected.
- Action tendencies regarding approach and avoidance lead to an immediate urge to act when exposed to a stimulus (Mogg et al. 2003). For example, a confrontation with a spider as a visual stimulus, a baby, or a tasty snack can cause a reaction pattern on the motoric level, causing an approach toward or a retreat from the stimulus. The same is true of other impulsive consumption goods.

These three layers of implicit cognition can provide complementary in-depth information about consumers' decision-making journey: Implicit attention can trigger an initial perception of a product or brand. Implicit associations can activate stored cognitions, thereby evoking an implicit attitude toward the consumption stimulus. Finally, approach tendencies can trigger an immediate urge to buy once an individual has been exposed to a desired product or brand.

1.2. Measurement of implicit cognitions

While researchers recognize the limitations of economic rationality widely, empirical research still struggles to open up to this refined, but methodologically more challenging, view of the drivers of human behaviour. Traditionally, quantitative marketing research has relied on explicit measures, such as traditional questionnaires, conjoint analyses, and discrete choice experiments to understand consumers' decision-making. These research methods and their underlying models of multi-attribute utility neglect automatic and unconscious cognitive processes, as well as their influence on decision-making.

Traditional survey techniques can hardly retrieve implicit cognition processes. Past memory traces, which themselves are not available for introspection or self-reporting, affect these processes (Greenwald and Banaji 1995). Consequently, reflective introspection in open-ended questions cannot capture these processes (Teichert and Heidel 2018). Social desirability can also contribute to biased answers (Brunel et al. 2004). This is especially true regarding assessing impulsive consumption decisions (King and Bruner 2000) during which ex-post rationalization camouflages the true triggers of decision-making.

Overcoming these limits, alternative methods measure cognitive processes either *indirectly* or *implicitly* (Greenwald and Banaji 1995). De Houwer et al. (2009, p. 347) define these methods as “measurement procedures that are caused in an automatic manner by psychological attributes.” The most prominent implicit cognition measurement paradigm is based on the phenomenon that individuals perform tasks congruent with their implicit cognitions faster. Tasks that are incompatible with consumers’ implicit cognitions take a bit longer, since cognitive resources need to be activated when there are no existing memory traces (Greenwald and Banaji 1995). An activating process causes a short inhibition period, which might be too subtle for human perception, but which can be unveiled with the help of computer-based technology sensitive enough to capture reactions in milliseconds. Consequently, prominent implicit measures compare relative response latencies measured in milliseconds to assess different types of implicit cognitive processes.

Distortions created by social desirability affect indirect measures less than direct self-report measures, because they do not depend on awareness of the reported process-

es (Heuer et al. 2007). Indirect measures do not require respondents to indicate their feelings about an object as they would do in an explicit measurement. Instead, respondents execute simple tasks, such as a categorization task, which neither informs them of what is assessed nor requests their self-report on it. Thereby, the stimuli spontaneously initiate answers without respondents being able to control their answering behaviour (Greenwald et al. 1998). Scholars have found that implicit measures are resistant to impression management: These measures might reveal traces of past experience that conflict with reachable cognition or that respondents reject explicitly (Nosek et al. 2005). These findings qualified initial applications in the social attitude research field on, for example, racial discrimination (Payne and Gawronski 2010).

Since then, various methods have evolved and been applied in different contexts. A recent overview (Znanewitz et al. 2018) reports on 20 implicit measurement methods. Among these, three prominent methodologies, which are widely used, stand out: the Stroop test for measuring implicit attention, the Implicit Association Test (IAT) for measuring implicit associations, and the Approach-Avoidance Task (AAT) for measuring implicit action tendencies.

In order to provide a broad overview of past research, we conducted a series of inquiries on the Web of Science. [1] Fig. 1 depicts how the scientific discourse on the three methods of implicit cognition measurement evolved over time. The figure shows that the number of publications within the broader field of behavioural decision-making (solid lines) has increased strongly. The Stroop tests lead historically and in their diffusion patterns, followed by the IAT tests. Approach and avoidance behaviour re-

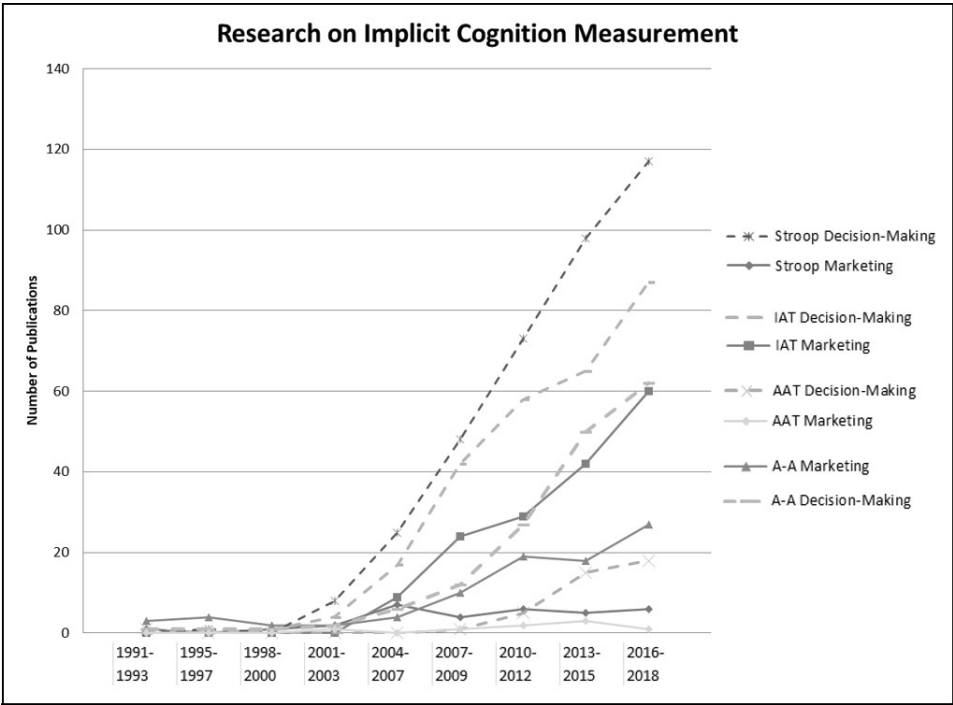


Fig. 1: Diffusion of implicit cognition measures (number of publications)

search began later and is not yet as widespread. Very few papers on approach and avoidance behaviour refer to the AAT test. Consequently, this measurement approach still lacks a shared research framework.

The graphs in *Fig. 1* that depict the diffusion into the marketing field show lags in time and quantity: This is most evident in respect of the Stroop test, because there are only a few marketing applications and they did not increase over time. The IAT measurement appears to be best established in the marketing field, but it still visibly lacks several years of diffusion. AAT measurements are nearly non-existent in marketing applications. However, the broader approach and avoidance concept is apparently gaining importance in marketing applications. Consequently, there is a remarkable and yet unexploited opportunity for implicit cognition measurement in marketing research.

1.3. Application Case

In order to illustrate how these implicit cognition measures are used in consumer behaviour research, we conducted a study on alcohol consumption in collaboration with the University of Witwatersrand (SA). The study's objective relates to social marketing, especially consumer protection. Alcohol consumption has characteristics common to different types of impulsive and compulsive consumption behaviours (Kyrios et al. 2004). Alcohol consumption is a potentially dysfunctional and common consumption practice among young adults (Bellis et al. 2009), especially in newly developing countries like South Africa. Owing to their social sensibility, standard questionnaires regularly fail to assess an alcohol consumption tendency and to reveal its underlying drivers. Alcohol consumption also relates to visceral and biological influences, making implicit measurements a necessity. We therefore applied an implicit cognition approach to study alcohol consumption tendencies among young South African adults. While identifying consumption patterns' (hidden) drivers generally helps marketers, our study aims at supporting an appropriate way for consumer protection agencies to address young adults.

We utilized the Ask your Brain (AYB) research platform as software for the implicit cognition measurement, which the first author developed. This non-profit software integrates the different implicit cognition measurements within a single online platform, thereby providing a shared user interface. Researchers can easily join the implicit cognition measures with traditional survey blocks (here implemented in Limesurvey) via automatic linkages. Surveys can be carried out on desktop computers and mobile devices. A user-friendly environment and a game-like graphic layout contribute to the survey's attractiveness for participants. Researchers can quickly configure experiments in a menu-driven backend according to a broad range of parameters (for details, see the morphological tables provided in each section). Subsequently, researchers can export all raw data as CSV files

and analyse them with multivariate software packages, such as STATA, SPSS or R.

In our application case, we join attentional bias, memory association, and action tendency measurements (see the following sections for details): Stroop task and Posner test measure attentional bias, the implicit association test (IAT) measures memory associations, and the Approach-Avoidance Task (AAT) measures action tendencies. The order of these tests is varied in four different routes with a Latin Square algorithm as outlined in *Tab. A1* (see appendix). Survey blocks, which were set up with the open source tool Limesurvey, interrupted the implicit measurement blocks. The researchers transferred an HTML parameter between the software packages, such that the survey questions were matched with all AYB cognition measures. Considerations of ethics did not allow providing students with alcohol. Consequently, the final test consisted of a discrete choice experiment: In 16 consecutive choice sets, consisting of four drink offers each, the participants had to choose between non-alcoholic and alcoholic beverages. The researchers calculated the percentage of alcoholic beverages chosen as a robust proxy of the respondent's alcohol consumption pattern. In addition, the researchers referred to the DCE for an ad hoc grouping: The participants were grouped into alcohol-inclined respondents (AIR) if they chose hard liquor (whiskey or vodka) at least once in the DCE, which consisted of eight choice sets; otherwise they were considered alcohol-disinclined respondents (ADR). Neither of the two measures has been validated and they are solely used for illustration purposes in the application case.

The researchers recruited students from two large universities in South Africa who participated in a computer room on campus or in the comfort of their own homes in exchange for a soft drink or course credit. Of the 225 participants who started the test, 104 completed the whole series, thereby resulting in a 46.2 % completion rate. The participants were, on average, 24.5 years old and 55.8 % were female. Based on the DCE choices obtained, the authors grouped the data according to alcohol-inclined ($n = 59$) and alcohol-disinclined respondents ($n = 45$).

2. Measuring Implicit Attentions

2.1. Research Paradigm

Attentional bias (Cox et al. 2006) constitutes a selective attention effect causing increased attention for an object. For instance, a person who had just eaten could easily miss a fast food chain's logo, but a fast food enthusiast craving a burger might recognize it quickly. A plethora of external stimuli bypasses the conscious mind, which only processes a small fraction of stimuli. Attentional bias filters the stimuli that are important for and relevant to the individual's goals, such as those that satisfy cravings for fatty food (Zanto et al. 2010). These filters also occur when different types of information are presented simultaneously.

Stroop (1935) measured reading interferences resulting from cognitive conflicts. The participants received cards containing a list of words for colours (100 words in total) printed in one of five different colours (e. g., the word “green” was printed in green, blue, red, brown, or purple). The control condition consisted of a list of colour squares. In this colour-naming task, naming the colour of the ink in which the words were printed entailed longer response times than the response times for the solid colour squares. This effect became known as the Stroop effect (MacLeod 1991) owing to a cognitive conflict while focusing on the target category (e. g., the colour of the ink). Adding a further information category (e. g., the colour of the word itself), prompts the participants to think about it, thereby inducing a cognitive conflict resulting in longer response times (Algom et al. 2004).

Later, researchers investigated the distraction caused by stimuli loaded emotionally via the emotional Stroop task (Williams et al. 1996). While the basic setting remains the same (presenting words in different colours and asking participants to respond to the displayed colour), the theoretical underpinnings and interpretations are different (Algom et al. 2004). In the original Stroop task, interference arises from two conflicting – but intertwined – sources of information. In contrast, the stimuli’s perceptual and semantic dimensions do not conflict in the Emotional Stroop task. Here, the argument is that emotional stimuli attract more visual attention than neutral stimuli and that this causes a longer response latency for the respondent (Cox et al. 2006; Field and Cox 2008).

Substituting emotional stimuli with addiction-related substances’ stimuli led to a rewording as “Addiction Stroop task” (Cox et al. 2006), whereby the interference’s scheme and interpretations are analogous to the Emotional Stroop task. Researchers apply this task widely in clinical psychological research to investigate selective attention with regard to addiction or mental disorders (Field and Cox 2008). For example, Waters and Feyerabend (2000) utilized the test to analyse smokers’ bias toward smoking-related stimuli. The addiction Stroop task was also used to compare attentional differences between light and heavy drinkers (Field et al. 2007) or to measure the effectiveness of drinkers’ alcohol attention-control (Fadardi and Cox 2009).

2.2. Applications in Marketing

As outlined above, researchers traditionally apply the Emotional Stroop task in clinical psychological research. More recently, the Emotional Stroop task also diffused into neuroscience applications, owing to a Stroop experiment’s easy setup and application in traditional laboratory settings and in brain scan environments. Researchers can combine the Stroop experiment with neuroscientific measures like EEG (Zanto et al. 2010) or fMRI (see e. g., Sachs et al. 2017). In contrast, marketing researchers typically conduct experiments outside of the laboratory. This might explain why Stroop experiments has not yet

diffused widely into marketing research, although initial studies demonstrate their usefulness to gain various insights in the realm of consumer behaviour research.

Extant Emotional Stroop studies in the field of marketing relate, first and foremost, to branding’s attentional effects. Keller et al. (2012) analysed the target group of children’s attentional biases toward food brands. They worked with children who were between seven and nine years old whom they grouped as either overweight or non-overweight according to their Body Mass Index (BMI). The children performed a modified food brand Stroop task that revealed an attentional bias. After completing the task, the children could eat a branded or unbranded test meal. The study showed how branding impacts on young children’s eating behaviour.

Scholars also used Emotional Stroop tests to assess brand relationship’s impact on attentional biases. Lee and Chiou (2013) used a modified Stroop task to analyse affiliation priming’s impact on the perception of social network brands. Very recently, academics investigated brand logo positioning’s impact on product packages via an Emotional Stroop task (Dong and Gleim 2018). The authors used the Emotional Stroop task to analyse differences between various brand logo positionings (high or low). They showed that consumers generally evaluate positive words, and particularly brand names, faster – measured by a shorter response time – when the positive words or brand names have a higher position on a product’s package. The study yielded insights about positioning brand names on products when familiarity with the brand is not taken into account.

Surprisingly few Emotional Stroop studies relate to advertising. In the sports marketing field, researchers investigated attentional bias for either exercise or sedentary lifestyle-related words (Berry 2006). Participants who did not exercise regularly showed an attentional bias toward sedentary-lifestyle words. The researchers provided conclusions for advertisement designs in the realm of health and exercise behaviour. Finally, the original Stroop test has been used to investigate the covariates of consumers’ decision-making: Researchers found that consumers who are in a negative mood have slower reaction times in a complex environment than participants who are in a positive mood (Braun-LaTour et al. 2007). This exemplifies the method’s usefulness to better understand consumers’ decision-making processes.

2.3. Experimental Setup

2.3.1. Basic Scheme

As stated, the original Stroop task has been repeatedly modified over the last decades. The Emotional Stroop task confronts respondents with words, supposedly emotional or neutral, which are printed in different colours. The respondents’ task is to name the displayed word’s colour as fast as possible. Hereby, researchers present a single stimulus for each trial and measure the response

latency per trial. Response inhibition manifests itself in relatively longer latencies, indicating an attentional interference in respect of emotionally laden stimuli.

Research applying the Stroop task does not follow a specific scheme, but varies in its experimental settings (see review in: Cox et al. 2006) and also in its applied scoring methods (see review in: Scarpina and Tagini 2017). The survey format and response modality range from cards to PC or mobile devices, with either vocal or manual input. In the usual unmasked condition, the word remains visible until the participant responds. In order to ensure that there are no reflective cognitive processes, few studies mask stimuli after a minimal display time (less than 100 ms). Emotional and neutral words can be either randomly intermixed or presented in blocks.

The length of Emotional Stroop tasks also vary significantly: For example, Fadardi and Cox (2009) used eight words and four colours, which they presented twice for each of the three different categories. This resulted in 64 trials per category and 192 trials in total. Munafò et al. (2003) conducted 180 trials in total (including 36 practice and buffer trials). The experiment of Steinhäuser and Hübner (2009), who conducted 960 trials with each participant, is an extreme case.

Algom et al. (2004) argue that the Emotional Stroop effect is list-based, since it does not only consider a specific item's interference, but the whole list of emotional stimuli applied in an experiment. Researchers must, thus, take particular care when selecting acceptable words (see review in: Cox et al. (2006), only selecting words that participants can read easily and understand intuitively. Short words with fewer than three syllables are preferred. The target population should use words that have a definite meaning frequently, but they should not be related to any particular colour. The neutrality of control words should be ensured. The number of stimuli should be identical across the word categories. In addition, the word length needs to match the frequency across categories to ensure the task's reliability (see e. g., McKenna and Sharma 1995). In order to individualize tests, the participants can provide pre-assessments through which the researchers can develop stimulus sets containing low-arousal and high-arousal words (see e. g., Frings et al. 2010).

2.3.2. Alternative Setups

Alternative test instruments investigate selective attention via a two-stage stimuli presentation process (see review in: Cox et al. 2006). In a visual probe task (Field and Cox 2008), two words or pictures initially appear briefly on the screen: A neutral object is positioned on one side of the screen, whereas the target object, for example, an emotional stimulus, is positioned on the other side of the screen. In a second step, a test stimulus (e. g., an upward-oriented or downward-oriented arrow) appears on one of the two sides of the screen and the respondent needs to react to it. Selective attention is diagnosed if the respondent reacts faster to the test stimulus

once it appears on the same side as the target object.

The Posner task (Posner 1980) also confronts the respondent with two simultaneously displayed stimuli, one emotional and one neutral stimulus, followed by a test stimulus (e. g., an arrow) to which the respondent has to react. Contrary to the test methods outlined above, the stimuli's display time is systematically varied, thereby making it possible to differentiate between two stages of attentional processes (Koster et al. 2006): facilitation (first 100 ms), as well as disengagement or avoidance (within 500 ms). Facilitation characterizes the ease with which attention can be directed toward an object. Disengagement describes the difficulties with diverting initial attention away from an object. In contrast, avoidance describes the tendency to divert attention away from an uncomfortable stimulus. Scholars found that the two stages of implicit attention evoke different responses: A very short presentation time (100 ms) can lead to stronger task engagement. In contrast, only longer presentation times triggered avoidance tendencies with regard to threatening pictures (Koster et al. 2006).

2.3.3. Online Implementation

Researchers usually perform Emotional Stroop experiments on a laboratory computer: They display words in different colours and the participants need to press the button on the keyboard that relates to the colour (Fadardi and Cox 2009; Kalanthroff et al. 2016). This setting is adapted to a more user-friendly environment for online and mobile surveys. AYB displays the Emotional Stroop task's word stimuli as graffiti on a wall and the targets are four paint buckets (Fig. 2). The respondent's task is to choose the paint bucket that matches the displayed colour. In an unmasked exposure condition, the word remains visible until the respondent clicks (in the case of a desktop PC) or touches (in the case of a smartphone) one of four paint buckets. In a masked condition, a pre-set time lapses, after which a mask, which is a random string of letters (e. g., STJXAMN) in the same colour, follows the stimulus word. Each mask has the same length as its preceding word.

The Posner task (Posner 1980) is implemented on two screens in AYB (Fig. 3). Experimenters can choose

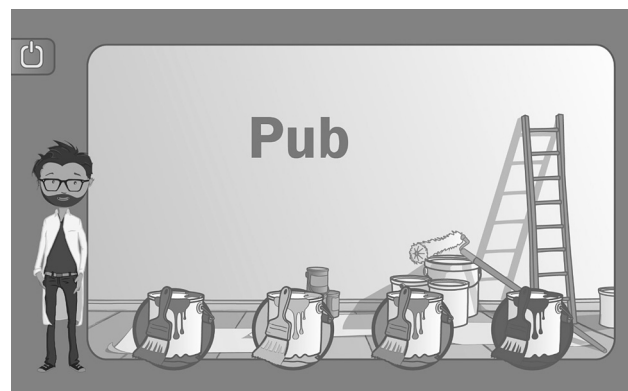


Fig. 2: Screenshot of AYB Emotional Stroop Task

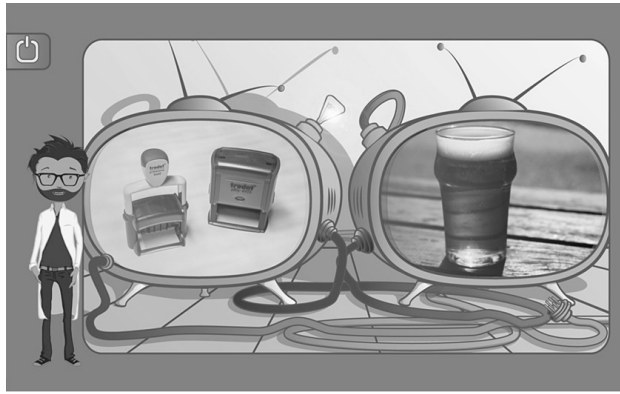


Fig. 3: Screenshot of AYB Posner Task

whether they should show stimuli only on one or on two screens simultaneously. In both cases, the participant needs to click or touch a target object that appears shortly after the stimulus disappears. In order to incentivize the respondents, they can receive feedback via the avatar; in addition, a score and an interpretation are also presented at the end of the experiment.

2.4. Data Analysis Steps

Researchers usually apply either an analysis of variance (ANOVA), a mixed ANOVA, or a repeated measures ANOVA as Stroop tasks' statistical analyses (see e. g., Braun-LaTour et al. 2007; Kalanthroff et al. 2016; Phillips et al. 2002; Waters and Feyerabend 2000). Reaction or task processing time is used as the dependent variable. The independent variables differ according to the experimental setup and research question.

Before data analysis, the researchers need to clean the data such that no extreme observations remain in the data set. The literature agrees that only correctly executed tasks should be used for analysis. However, there are various methods, differing in their degree of sophistication, for identifying extreme observations. A legit cutoff is frequently defined for the reaction time, which can, according to Ratcliff (1993), vary, for example, owing to task difficulty. The researcher has to determine the cutoff, which leads to the reaction time exclusions below and above a certain threshold. As a rule of thumb, academics usually understand response times below 200 ms or above 2000–3000 ms as outliers (Munafò et al. 2003; Waters et al. 2005). In a statistically more sophisticated approach, Schmiedek et al. (2007) used a quantile maximum probability estimator for identifying outliers in a reaction time environment by means of a recursive process. In a first step, reaction times under 200 ms and larger than four standard deviations above the individual's mean reaction time are excluded from the analysis. Thereafter, the procedure is repeated until the response data contain no outliers. As an example, Steinhäuser and Hübner (2009) apply the method for a Stroop task. This method's advantage is that it achieves an approximated normal distribution better than by applying standardized cutoff values. However, this method obstructs a direct

comparison of the estimated effect sizes across different studies.

The number of identified outliers can be quite large. This number also varies largely across studies: For example, Hornung et al. (2017) discarded 12.8 % of their data. Steinhäuser and Hübner (2009), who also implemented the method of Schmiedek et al. (2007), discarded less than 1 % of their data.

After discarding the outliers, the researchers calculate the mean reaction times of each category (emotional vs. neutral) on the individual respondent level. Emotional Stroop interference scores are also widely used as a robust base for comparison (see Cox et al. 2006 for an application to alcohol-related stimuli). In its basic form, we define this measure as the difference between the individual's mean reaction time (RT) to the experimental category's stimuli (e. g., emotional) and to the neutral category's stimuli.

$$\text{Interference} = \text{Mean}(RT_{\text{exp}}) - \text{Mean}(RT_{\text{neutral}}) \quad (1)$$

ANOVA analyses evaluate the mean reaction times or interference scores' differences between different experimental groups. Researchers find an Emotional Stroop effect via a significant difference between the mean response times in trials that include emotional words and trials comprising neutral words (Algom et al. 2004). Error rates can also be used as dependent variables (see e. g., van Holst et al. 2012). The error rates provide insights into emotional or addiction-related words' semantic processing. Furthermore, researchers can use effect sizes like Cohen's *d*, which is specified as the mean interference scores' difference between two groups divided by its pooled standard deviations, or η^2 , to show the variance that a specific independent grouping variable explains in the model. For models with more than one independent variable, it is appropriate to use η^2 .

2.5. Methodological summary and limitations

Tab. 1 summarizes the methodological conventions applied to Stroop experiments (also see the meta-analysis of Cox et al. 2006). In offline settings, the experimenter can present a list of words, which must be spoken aloud. The more common approach is to present individual stimuli on a PC, while two stimuli categories are mainly applied in Emotional Stroop tasks. Researchers can decide to program their task manually by using basic freeware like R; they can also use commercial software like Inquisit or E-Prime. The nonprofit software, AYB, does not require any programming, owing to its graphical user interface. The level of data analysis varies. In traditional Stroop experiments, researchers calculate overall reaction times or Stroop interference scores to analyse the data. In Emotional Stroop tasks, each trial's reaction times are measured and aggregated for each individual, which forms the basis for calculating Stroop interference scores.

Although Stroop tasks are easily set up and the data analysis steps are straightforward, the method also has cer-

Configuration	Environment	Online*		Laboratory*	
	Input device	PC/Keyboard	PC/Mouse*	Mobile/Touchscreen*	Word naming aloud
	Software	Basic Freeware		Nonprofit Software*	Commercial Software
Study Setup	Stimuli categories	Two (neutral vs. emotional)*		Three (congruent vs. incongruent vs. neutral)*	
	Stimuli types	Words in different colors*		Pictures overlaid with words	
	Stimuli selection	Word frequency	Word length	Established words	Word-rating by participants
	Number of stimuli/category	Low (< 5 stimuli)	Medium (< 15 stimuli)	High (≥ 15 stimuli)	
	Test length	Brief (< 100 trials)*		Regular (< 200 trials)*	Long (≥ 200 trials)*
Data treatment	Practice trials	Exclude first block of trials*		Exclude first trial(s) in each block*	No exclusion of any trials
	Error treatment	No treatment		Exclude errors	
	Extreme observations	No treatment		Substitute with cutoff value	Exclude observations
	Differentiation lower & upper limit	Exclude observations below threshold value (e. g. 200 ms)		Exclude observations above threshold value (e. g. 3000 ms)	Exclude observations beyond four-times standard deviation above individual mean
Data analysis (standard software)	Dependent variable	Stroop interference score	Reaction time	Total processing time	Errors
	Aggregation level for DV	Population level	Individual level	Single Stimuli level	Category level
	Analysis method	ANOVA		Mixed ANOVA	Repeated measures ANOVA
	Reliability checks	Test-Retest-Reliability		Exchange of stimuli	Learning curve effects

Tab. 1: Methodological conventions of Stroop Tests: Traditional default in grey, AYB study example in italics, *) Available options within core AYB program

tain critical issues. In his early research, Stroop (1935) noted that participants exhibit a learning curve while performing the task, especially when the task takes longer. Long and Prat (2002) showed that high capacity individuals can also evoke lower interferences. These limitations also prevail in the Emotional Stroop task, whereby habituation reduces the interference during the experiment (McKenna and Sharma 1995). Exchanging stimuli words between the experimental blocks can help maintain a persisting interference toward emotional stimuli. Larsen et al. (2006) provide a review of how lexical characteristics can influence the response latencies. Emotional words are, for example, more often longer in length than control stimuli, which can also contribute to delayed latencies. Consequently, researchers should be very precise when choosing appropriate stimuli.

Researchers have tested test-retest reliability in many experiments. Past experiments showed that the Stroop task's test-retest reliability is high when mean response times are used as the dependent variable (see e. g., Siegrist 1997; Strauss et al. 2005). In contrast, interference scores were found unreliable with regard to, for example, taboo language or words related to anger in Emotional

Stroop tasks. Only non-clinical respondents participated in the experiments.

2.6. Case Illustration

Generally, researchers apply Stroop tasks to compare stimuli effects between different groups of participants. Our exemplary individual stimulus for each trial Emotional Stroop task analyses whether attentional biases differ between alcohol-inclined respondents (AIR) and alcohol-disinclined respondents (ADR), depending on the stimuli displayed, that is, alcohol vs. control stimuli. Following an experimental set-up by Stormark et al. (1997), we compared alcohol-related stimuli (e. g., "beer," "pub," "vodka") to neutral unit-related stimuli (e. g., "meter," "gram," "week"). [2]

Each participant conducted 78 trials in the Emotional Stroop task. These trials were split into seven blocks of roughly the same length to counteract fatigue effects in the respondents. Neutral and alcohol-related stimuli varied randomly within each block. We excluded the experiment's first block from the analysis, because it served as practice for the participants to adjust to the task. Each

block’s first trial was also excluded from the analysis, because the participants might have re-oriented themselves at the beginning of each block. In total, the 104 participants conducted 6,864 trials in this part of the experiment.

We cleaned the data by removing outliers, following the Schmiedek et al. (2007) procedure. For each respondent, we calculated the means and standard deviations of correct responses with reaction times faster than 200 ms. Trials that took longer than four times the standard deviation above the mean were then excluded. The procedure was repeated until no outliers remained in the respondents’ data. We excluded only 80 observations across all the respondents from the analysis, because their responses were either false or identified outliers (~1.1 %). The amount of discarded data is lower than those in a variety of Stroop studies, thereby indicating comparatively low interferences.

The Stroop interference (see Fadardi and Cox 2009) for alcohol stimuli is calculated by subtracting the respondents’ mean reaction time (latency) to neutral stimuli from that to alcohol-related stimuli. In this instance, researchers expect the alcohol-inclined respondents to show a positive effect, revealing an attentional bias toward the alcohol-related stimuli. However, as Tab. 2 shows, the findings do not support this hypothesized effect. Neither of the two groups exhibits a significant attentional bias toward or against alcohol-related stimuli. Furthermore, there is no significant difference between the two groups, $F(1,102) = .05, p = .82, \eta^2 = .000$.

Based on the above findings, we cannot replicate previous findings showing heavy alcohol drinkers’ significantly slower reactions to alcohol stimuli in the Stroop task (see Fadardi and Cox 2009; Lusher et al. 2004). The

relatively small sample size and the comparatively homogeneous student population with no expected cases of alcohol addiction could clarify this inability. We, therefore, repeated the analysis in a more narrowly defined setting by only using “Vodka” and “Liquor” as words representing the experimental category. In our opinion, the attentional effects toward hard liquor are more likely to differ between alcohol-inclined and disinclined persons than between the attentional effects toward casual alcoholic beverages.

Tab. 3 shows that the Stroop interference score’s mean value becomes positive (even if it is not significant) for the AIR group, thereby indicating a weak attentional bias toward their “hard liquor” stimuli. Furthermore, the two groups somehow exhibit different attentional biases toward hard liquor, with the AIR group showing longer average latencies (950 ms) than the ADR group (868 ms). However, with $F(1,102) = 2.04, p = .156, \eta^2 = .02$, this difference between the two groups is still not significant.

Tab. 2 and Tab. 3 show that alcohol-inclined respondents’ responses took longer than those of the alcohol-disinclined respondents. The alcohol-related stimuli’s repeated presentation might have caused alcohol drinkers to exhibit an overall response inhibition. This can be regarded as the displayed alcohol stimuli’s longer-lasting carry-over effect on the overall stimulus response (Waters et al. 2005).

In order to check for this possible explanation, we divided the dataset into initial blocks (2–4) and final blocks (5–7). Fig. 4 shows that the response latency difference between the two groups becomes particularly evident in the last three blocks: Whereas the ADR group becomes faster in the last three blocks, the AIR’s response laten-

	ADR (<i>n</i> = 45)	AIR (<i>n</i> = 59)	Total (<i>n</i> = 104)
Latency for alc stimuli	878.98 ± 60.26 (4.58)	935.27 ± 52.63 (3.50)	910.91 ± 406.01 (20.31)
Latency for unit stimuli	875.58 ± 62.79 (4.77)	929.43 ± 54.84 (3.64)	906.13 ± 422.45 (21.13)
Stroop interference score	3.40 ± 16.15 (1.23)	5.84 ± 14.10 (.94)	4.78 ± 107.81 (5.39)

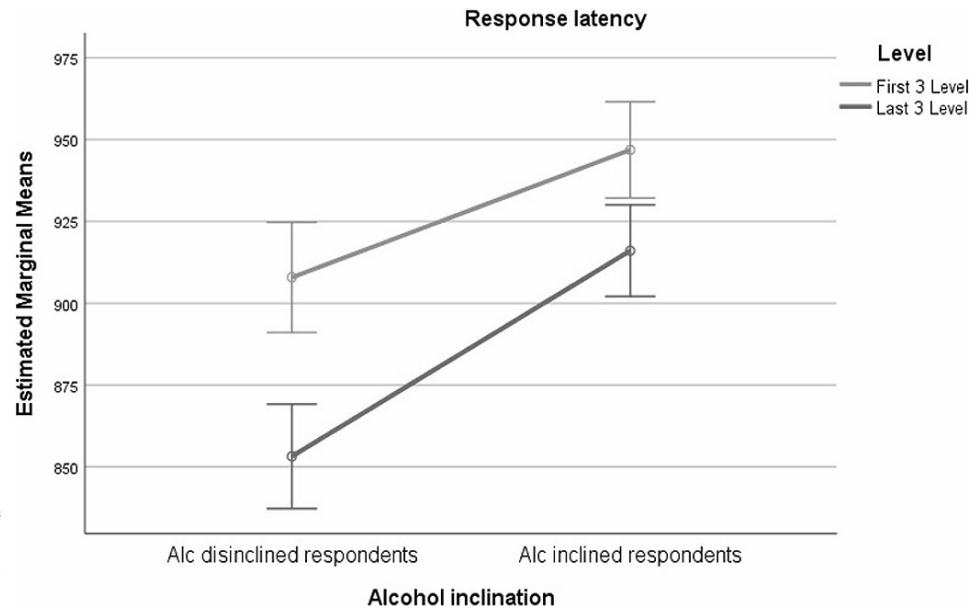
Notes: Mean values ± 95 % confidence intervals, (standard error of the means) of alcohol-inclined (AIR) and alcohol-disinclined respondents (ADR)

Tab. 2: Response latencies and Stroop interference score for all alcohol-related stimuli

	ADR (<i>n</i> = 45)	AIR (<i>n</i> = 59)	Total (<i>n</i> = 104)
Latency for alc stimuli	867.91 ± 65.81 (33.58)	949.66 ± 57.48 (29.33)	914.29 ± 446.53 (227.82)
Latency for unit stimuli	875.58 ± 62.79 (32.04)	929.43 ± 54.84 (27.98)	906.13 ± 422.45 (215.54)
Stroop interference score	-7.67 ± 28.82 (14.70)	20.23 ± 25.17 (12.84)	8.16 ± 194.31 (99.14)

Notes: Mean values ± 95 % confidence intervals, (standard error of the means) of alcohol-inclined (AIR) and alcohol-disinclined respondents (ADR)

Tab. 3: Response latencies and Stroop interference score for “Vodka” and “Liquor”



Notes: Error bars: 95 % confidence interval

Fig. 4: Stroop response time differences between ADR and AIR

cies do not improve a lot. The overall interaction effect between alcohol inclination and the test block is, however, insignificant, $F(1,5925) = 2.35$, $p = .13$, $\eta^2 = .000$. Consistent with McKenna and Sharma's (1995) findings, the participants overcome overall interference when the experiment proceeds. However, the AIR group's response times did not improve much. This shows that the alcohol stimuli's emotionality persisted longer in their case than in the other group.

While this is merely weak evidence of the AIR's attentional bias, it should be kept in mind that the student sample probably did not include alcohol-dependent respondents. The findings can therefore be regarded as an initial – though limited – proof that this method is applicable in the context of regular consumption patterns.

In order to further investigate the predictive potential for real choice situations, we analysed whether either the Stroop interference score or the participants' mean reaction times influence the percentage of alcoholic drinks chosen later in the DCE. Tab. 4 shows that no significant effect was found in the data. Consequently, we cannot find an influence of the Stroop interference score on the choice of alcoholic drinks at the individual level. The results do not change if we only consider "Vodka" and "Liquor" as alcoholic stimuli.

Model	B	S.E.	t	p	R ²
Stroop interference	-.0001	.0005	-.22	.825	.001
Mean reaction time	.0001	.0001	.88	.381	.008

Notes: Dependent variable: Percentage of alcoholic drinks chosen. S.E. = Standard Error

Tab. 4: Regression coefficients for Stroop interference score and mean reaction time

3. Measuring Implicit Associations

3.1. Research Paradigm

Implicit associations form a second layer of implicit cognitions. This concept refers to mental representations automatically activated by means of spreading activation even when a stimulus is simply presented (Anderson 1983). For example, mentioning the word "dog" may trigger spontaneous positive associations like "cute" or "lovely" if a respondent grew up with such an animal. Alternatively, it may trigger negative associations like "danger" or "enemy" if a dog had previously attacked the respondent. Such associations and the resulting emotions are not consciously reflected, but are the result of a learning process. They lead to implicit attitudes, which are valenced associations with an object (Fazio and Olson 2003). Even if one explicitly resists a connection between concepts – for example, dog – bad – this implicit tendency may alter behaviour, especially if one lacks the motivation or resources to override it (Friesen et al. 2006).

The Implicit Association Test or IAT (Greenwald et al. 1998) measures the differential associations between (two) target concepts and (two) attribute dimensions. The IAT is the standard instrument for measuring implicit attitudes. Researchers measure cognitive content indirectly via reaction latencies that occur when a respondent performs a task. Scholars particularly compare responses from combinations of a bipolar target category and a bipolar attribute category. Category combinations congruent with the respondent's intuition (compatible) facilitate task performance and, thus, result in shorter reaction times. Incompatible combinations, in contrast, cause interferences that result in longer latencies.

Referring to the above example, stimuli from the target category "dogs" and a contrasting category "cats" would be presented together with either positive or negative at-

tributes on a PC screen, requesting a key stroke as response. If respondents associate dogs more strongly with positive attributes than cats, they would respond more quickly to the stimulus paired accordingly. The IAT score calculation builds on this difference between the relative response times.

The IAT is by far the single most-used test procedure in the implicit cognition realm. The good psychometric properties of the IAT, as well as its standardized procedure, support its extensive use. IAT results are, however, surprisingly unresponsive to procedural variations (Nosek et al. 2005). The IAT can explain variance in a number of different behaviours. It specifically outperforms explicit attitude measures in social sensitive studies, for example, on interracial and intergroup behaviour (Greenwald et al. 2009).

Greenwald and colleagues published their seminal IAT article in 1998. Since then, several review articles and meta-analyses have contributed to understanding the IAT. Fazio and Olson (2003), as well as Banaji (2013), wrote noteworthy review articles. De Houwer et al. (2009) provide a relatively normative conceptual framework for implicit measures. Hofmann et al. (2005) conducted a meta-analysis on the correlation between explicit self-reports and IAT measures. They found that correlations increase with the respondents' increasing spontaneity and the conceptual correspondence between implicit and explicit measures. Greenwald et al. (2009) performed a meta-analysis on the IAT's predictive validity. They found that implicit and explicit measures can be effective predictors of actual behaviour. However, explicit attitude measures exhibit far greater variability, while predictive validity is impaired for socially sensitive topics.

3.2. Applications in Marketing

Of all the implicit cognition measures, the IAT has – by far – diffused the most into the marketing domain. In the consumer behaviour realm, Dimofte (2010) provides a broad review of implicit cognition with a strong focus on the IAT. Gregg and Klymowsky's review (2013) summarizes the IAT applications, particularly in consumer behaviour research. The following overview (Tab. 5) summarizes the IAT applications along six categories: brand attitudes, social groups, preference formation, cue associations/metaphors, and sustainable consumption.

Early research focused on brand attitudes (Brunel et al. 2004; Maison et al. 2004), for example, attitude comparisons between Coca-Cola vs. Pepsi (Maison et al. 2004) and Apple vs. Windows (Brunel et al. 2004). Implicit associations play a major role in marketing research, because brands evoke an image that is characterised by brand associations imprinted on the consumer's memory. However, the IAT's additional value is rather limited: Consumers gladly disclose their brand attitudes. Hence, the IAT results on brand attitudes usually exhibit high correlations with explicit measures (Hofmann et al. 2005).

Implicit associations with social groups are also of interest to marketing research. For example, Brunel et al. (2004) investigated racial preferences regarding spokespersons in advertisements. Luna et al. (2008) show that cueing bicultural persons with a particular language causes them to switch between concepts' frames, such as masculinity/femininity.

IAT applications can inform Marketing theory about the cognitive mechanisms of *preference formation*. For example, Fukawa and Niedrich (2015) use the IAT as a process measure to show that supraliminal primes can affect product preference via conceptual fluency. Forehand and Perkins (2005) demonstrate that consumers consciously

Brand	Social groups	Preference formation processes
Angle and Forehand (2016), IJRM	Angle et al. (2017), JCP	Chang and Ko (2016), P&M
Dagogo-Jack and Forehand (2018), JMR	Brunel et al. (2004), JCP	Chaxel and Han (2018), JCP
Dimofte and Yalch (2007), JCP	Forehand et al. (2011), JCP	Czellar and Luna (2010), JCP
Dimofte and Yalch (2006), JCR	Humphreys and Latour (2013), JCR	Forehand and Perkins (2005), JCR
Gibson (2008), JCR	Luna et al. (2008), JCR	Fukawa and Niedrich (2015), P&M
Maison et al. (2004), JCP	Nevid and McClelland (2010), P&M	Naylor et al. (2006), JCP
Nevid and Pastva (2014), JCR		Scarabis et al. (2006), P&M
Sprott et al. (2009), JMR		
Cue associations / Metaphors	Sustainable marketing	Impulsive consumption
Cian et al. (2015), JCR	Brough et al. (2016), JCP	Friese et al. (2006), P&M
Rozin et al. (2012), JCR	Lee et al. (2018), P&M	May and Irmak (2018), JCR
Yan (2016), JCR	Luchs et al. (2010), JM	Raghunathan et al. (2006), JM
Yan and Pena-Marin (2017), JCR	Zlatevska and Spence (2012), P&M	
Lee et al. (2014), JCR		

Notes: IJRM = International Journal of Research in Marketing, JCR = Journal of Consumer Research, JCP = Journal of Consumer Psychology, JM = Journal of Marketing, JMR = Journal of Marketing Research, P&M = Psychology & Marketing

Tab. 5: IAT studies published in top-tier marketing journals

adjust positive image transfer from celebrity voice-overs if they consciously recognize the voice.

In order to arrive at a satisfactory decision, consumers have to compute a plethora of information. Simplifying the understanding of abstract concepts, consumers often rely on *cue associations* and *conceptual metaphors* that draw on concrete experiences from the physical world (Cian et al. 2015). Using an IAT (among other measures), Cian et al. (2015) show, for example, that individuals map the concepts of rationality and emotion according to the vertical difference between the head and the heart. Rozin et al. (2012) demonstrate that consumers metaphorically relate the concept of meat to masculinity. Yan (2016) posits that precise numbers are considered more masculine than round numbers.

Cue associations can determine attitude formation in a wide range of applications. Green, eco-friendly, or otherwise ethical products and *sustainable consumption* have recently gained particular traction. For example, Brough et al. (2016) show that eco-friendly products are commonly associated with femininity, which poses a potential identity threat for male consumers. Luchs et al. (2010) show that sustainability can cause negative consumer reactions when valuing strength-related attributes. Lee et al. (2018) demonstrate that consumers associate “greenwashed products” with “self-interested companies.” Recently, Ackermann et al. (2018) applied the IAT in an innovation context where attitudes toward new product categories are formed.

Impulsive consumption is another application field for IAT measures. Explicit measures reflect more deliberative behaviour and implicit measures reflect rather spontaneous behaviour. Frieze et al. (2006) show, for example, that individuals follow their implicit tendencies under time pressure and otherwise act according to their ex-

PLICIT attitudes. Florack et al. (2010) argue that impulsive preferences are more likely to prevail if individuals focus on acquiring a desired state (promotion focus) than if they avoid an undesired state (prevention focus).

3.3. Experimental Setup

3.3.1. Basic Scheme

Greenwald et al. (1998) defined a standardized setup in five blocks (see Tab. 6). In the first block, two object categories (e. g., cats and dogs) are allocated to one response key each. The participant is shown either a word or picture (e. g., shepherd), which he or she must then assign as quickly as possible to one of the two categories by pressing the designated key. In the second block, the same procedure is repeated with association categories (e. g., positive and negative words) replacing the object categories. A target and an attribute (e. g., dogs and positive; cats and negative) are matched in the third block. The randomly presented stimuli belong to the target category or the attribute category and the respondent has to decide on the category. The remaining two blocks repeat blocks 1 and 3 in a reverse display setting so that the target and attribute categories are counter-matched (e. g., cats and positive; dogs and negative). Only blocks 3 and 5 are used for estimation purposes. The differences in the average latencies of blocks 3 and 5, which reversely match the targets and the attributes, are calculated to obtain the standardized D-Score (see below).

3.3.2. Alternative Setups

A number of measurement procedures are closely related to the IAT. A few of these procedures measure similar constructs, while others address methodological shortcomings by varying the procedure. Tab. 7 provides an overview of the IAT’s prominent alternatives and variations.

Tab. 6: Standard IAT Procedure (based on Greenwald et al. 1998)

Block #	Block used for	Left key	Right key
1	Practice	Target 1	Target 2
2	Practice	Attribute 1	Attribute 2
3	Test	Target 1 + Attribute 1	Target 2 + Attribute 2
4	Practice	Target 2	Target 1
5	Test	Target 2 + Attribute 1	Target 1 + Attribute 2

Tab. 7: Overview of prominent IAT alternatives

Procedure	Mechanism	Recoding-Free	Single Object	Multiple Objects
Affective Priming	Priming		X	X
Affect Misattribution Procedure	Priming		X	X
Brief IAT	Categorization			
Extrinsic Affective Simon Task	Simon Effect	X	X	X
Go/No-Go Association Task	Signal detection	X	X	X
Recoding-Free IAT	Categorization	X		
Single-Category IAT	Categorization		X	

Prior to the IAT, researchers used other procedures as unobtrusive measures of attitudes, i. e., valenced associations with an object. Fazio et al. (1995) introduced Affective Priming (AP) as an attitude measure. This procedure relies on the phenomenon that a stimulus presented at an earlier stage can affect individuals’ task performance. The participants are shown a stimulus and later an adjective, which they have to categorize as, for example, “negative” or “positive.” A quick response is facilitated if the preceding stimulus is evaluation congruent (e. g., a picture of an unpopular animal followed by the word “disgusting”), which results in shorter reaction times. The opposite is true, if the stimulus and the adjective are incongruent. A similar logic is applied in the Affect Misattribution Procedure (AMP), wherein an affect-laden prime is presented to the participants followed by a target object, which they must classify as relatively pleasant or unpleasant. The results show that the participants misattribute the preceding prime’s valence to the subsequent stimulus, even when they are warned that the preceding stimulus might affect their response (Payne et al. 2005).

Numerous variants of the IAT emerged to address its potential methodological issues (Tab. 7). For example, individuals might exploit similarities between objects that are not part of the association task – in other words, they recode the task. As a remedy, Rothermund et al. (2009) proposed the Recoding-Free IAT (IAT-RF), whereby, the compatible and incompatible assignments are not divided into separate blocks, but randomly switched between trials. The Extrinsic Affective Simon Task (EAST, De Houwer 2003) offers another way of preventing recoding: Here the participants are asked to categorize white words based on stimulus valence and coloured words based on colour. The Simon effect (Simon 1969) causes that the color of a valenced word (e. g., “flowers”) is classified faster, if the response key for colour matches the valence key.

While the classic IAT setup requires two target concepts, the EAST, the Go/No-Go Association Task (GNAT; Nosek and Banaji 2001), and the Single-Category IAT (SC-IAT; Karpinski and Steinman 2006), present viable alternatives if a single attitude object is measured. The GNAT compels the participant to press a single key (“Go”) before a response deadline, if the presented word belongs either to the target or to the attribute category, which is permanently shown on the screen. If the stimulus does not belong to the target or the attribute category – that is, if it represents noise – the respondent should not press the key (“No Go”). Instead of reaction times, the GNAT utilizes sensitivity to distinguish between signal and noise as a measure for an association between the target and the attribute category. The SC-IAT is similar to the classic IAT procedure. However, the SC-IAT omits the second target and combines the object with each key that is associated with an attribute (see Tab. 8).

Finally, the Brief IAT (BIAT; Sriram and Greenwald 2009) utilizes the classic IAT’s block setup, but only a

Blocks	Left Key	Right Key
Practice	Attribute 1 + Target	Attribute 2
Test	Attribute 1 + Target	Attribute 2
Practice	Attribute 1	Attribute 2 + Target
Test	Attribute 1	Attribute 2 + Target

Tab. 8: Procedure of Single Category-IAT

third of the classic IAT’s trials. The authors show that even this reduced setup can provide satisfactory validity and similar results.

3.3.3. Online Implementation

AYB replicates the IAT as a card game. The participant virtually receives a card that he must slide to the left card stack or to the right card stack. This modifies the original IAT’s task based on the input of two keys on a keyboard. The sliding task enables mouse trajectory tracking as an additional measure of the decision-making processes, which is often realized with additional software like Mouse Tracker (Freeman and Ambady 2010). For example, Yu et al. (2012) show that, although the participants responded mostly correctly, during incompatible trials, their mouse movements were continuously drawn to the alternative response. The AYB platform captures and timestamps the mouse movement coordinates without installing another software package (see Fig. 5).

3.4. Data Analysis Steps

The IAT’s statistical analyses build on a standardized difference score between mean latencies in compatible and incompatible situations, also referred to as the D-Score. The mean reaction times obtained in blocks 3 and 5 of the standard IAT (Tab. 6) are subtracted from each other and divided by the respondent’s pooled standard deviation (Greenwald et al. 2003). The D-Score mirrors the implicit association’s magnitude and direction. The following equation shows the D-Score computation’s logic:

$$D = \frac{M(incompatible) - M(compatible)}{Pooled\ SD} \tag{2}$$

The D-Score increases if a large interference occurs, i.e. a longer mean reaction time during the incompatible

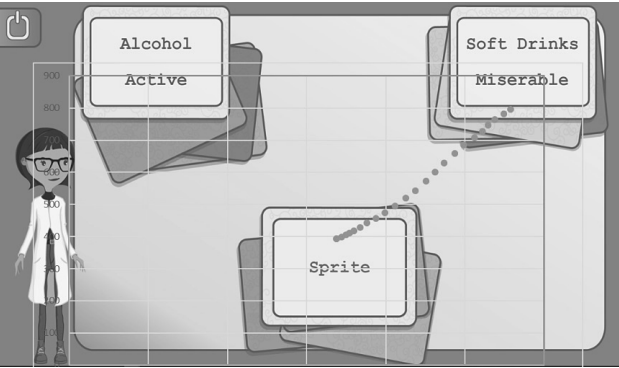


Fig. 5: Screenshot of AYB IAT with mouse trajectories displayed

block. D also increases if a facilitation effect occurs, owing to the already familiar combination of concepts, i.e. in case of a shorter reaction time during the compatible block. The Pooled Standard Deviation (SD) of both blocks in the denominator serves as a robust correction measure. If a high amount of variation occurs in response latencies, the mean difference probably results from chance, and, therefore, D decreases. Conversely, a rather small variation indicates that the mean difference is not produced by chance, and, therefore, D increases.

Importantly, when specifying the D-Score, the researcher must specify *ex ante* the category combinations' compatibility or incompatibility. In certain cases, the compatible block is obvious; for example, most individuals should associate "flowers" with "positive" and "insects" with "negative." In other cases, the denomination depends on the formulated hypotheses or on documentation convenience. For example, the application case below investigates whether individuals who associate "Alcohol" and "Active" choose a higher number of alcoholic drinks. Defining the reverse combination ("Alcohol" and "Miserable") as compatible only changes the coefficient's sign, but demands slightly more effort to interpret the results.

There are multiple conventions for computing the D-Score. The most prominent methods are the conventional algorithm Greenwald et al. (1998) proposed in their seminar article as well as their improved algorithm (Greenwald et al. 2003). The two algorithms differ from each other with regard to the considered blocks, the way extreme observations (at the trial level and the subject level), as well as incorrect trials and latency transformations are treated. Most importantly, the conventional algorithm applies the logarithm, while the improved algorithm builds on untransformed latencies. *Tab. 9* compares the two algorithms. The conventional algorithm treats the data more conservatively, while the improved algorithm makes the most of the existing data with fewer considerations regarding the latency distributions and the potential effects of extreme observations.

Researchers generally utilize Cronbach's alpha to assess the D-Scores' reliability. Furthermore, researchers can check internal consistency by randomly splitting the respondents' observations and observing how the two computed D-Scores correspond (split-half reliability). A Spearman-Brown correction is often applied afterwards (e. g. De Houwer and Bruycker 2007). The test-retest re-

liability is another form of reliability. Here, the test is administered a second time at a later stage. Studies show that the IAT results are not state dependent, since the results barely change in the test-retest settings. The reported coefficients range from about 0.7 for test-retest reliability (Bosson et al. 2000) to about 0.9 for split-half reliability (Greenwald and Nosek 2001), exceeding the reliability coefficient observed in other implicit cognition measures.

Mouse trajectories offer another analysis opportunity. Maximum deviations, motor trajectories, as well as velocity and acceleration profiles, offer additional insights that go beyond reaction times. To convert the mouse coordinates into the Euclidean distance, which represents the distance between the cursor and the response option (Hehman et al. 2015), the following equation can be applied:

$$distance((x,y),(a,b)) = \sqrt{(x-a)^2 + (y-b)^2} \tag{3}$$

with (x, y) being any location and (a, b) being either of the response options. To adjust the distance, proportional proximity is often calculated as $1 - distance / max(distance)$.

The mouse trajectories can take two basic forms: smooth or abrupt. The latter is redirected in midflight, while the former follows its initial direction (Hehman et al. 2015). Freeman (2014) formally describes abrupt shifts as those with a maximum deviation coefficient that exceeds 0.9, which signifies an initial attraction toward the unselected category. While spatial data analyses can already be insightful, researchers can also consider timestamped data to analyse temporal dynamics. Researchers can specifically use acceleration and velocity profile analyses to determine at which moment a participant commits to his response (Hehman et al. 2015).

3.5. Methodological summary and limitations

Researchers have applied the IAT extensively in diverse settings, which has led to a myriad of procedural variations. *Tab. 10* provides an overview of alternatives for the study setup, data treatment, and data analysis. The following section guides the reader through these options, enabling the IAT researchers to select the most appropriate solution.

In a preparatory phase, the researchers must decide in which environment they want to conduct the experiment, which input device and software they will use, how

Tab. 9: Comparison of prominent D-Score algorithms

Treatment	Conventional algorithm	Improved algorithm
Blocks	Only long blocks	All; two scores are created and averaged
Trials	Drop the first two trials of each block; recode latencies outside 300/3000 ms	Eliminate trials > 10.000 ms
Subject	Nonsystematic elimination	Eliminate subjects with > 10 % of trials < 300 ms
Incorrect trials	No treatment	Replace error latencies with block mean + 600 ms
Transformation	Log-transform	No transformation, use mean and SD

Data preparation	Environment	Online*		Laboratory	
	Input device	PC/Keyboard*	PC/Mouse*	Mobile/ Touchscreen*	Word spoken aloud
	Software	Basic freeware		Nonprofit software*	Commercial software
Study setup	Stimuli categories	Single category*		Two target categories*	Multiple target categories (>2)
	Stimuli types	Words*		Pictures*	
	Stimuli selection	Word frequency	Word length	Established words	Word-rating by participants
	Test length	Brief (<80 trials)*		Regular (180 trials)*	Long (>180 trials)*
Data treatment	Practice trials	Exclude practice block		Exclude first two tasks in each block	No exclusion of trials
	Error treatment	No treatment	Exclude errors	Recode errors	Separate
	Extreme observations	No treatment		Trimming	Winsorizing
	Differentiation lower & upper limit	Recode observations < 300 ms		Recode observations > 3000 ms	Exclude observations > 10.000 ms
Data analysis (standard software)	Dependent variable	D-Score	G-Score, Worst performance rule	Reaction time	Errors Mouse move-ment path
	Aggregation level for DV	Population mean values		Individual mean values	Block level Trial level
	Analysis method	ANOVA / Repeated ANOVA		Hierarchical regres-sion; mixed models	Robust statistics QUAD model
	Reliability checks	Internal consistency		Test-retest-reliability	Validity Predictive validity

Tab. 10: Methodological conventions of IAT: Traditional default in grey, AYB Study example in italics, *) Available options within core AYB program

many categories the test will comprise, and how long the test should be. While laboratory experiments can control for external influences, online experiments provide more ecological validity. In the future, GeoPosition and time might offer additional information that researchers can take into account during online field studies. While data collected outside the laboratory is potentially contaminated with noise, Brown et al. (2014) contend that a large sample size can compensate for this. While the default input mechanism is the standard computer and computer keyboard, mouse input is needed to obtain additional information via a mouse path analysis (see e. g., Yu et al. 2012). Touchscreen mobile devices are a promising alternative when executing online field experiments.

Regarding software, researchers can choose between commercial offers (e. g., Millisecond Inquisit, E-Prime), free open-source options (OpenIAT, PsychoPy), and non-profit software (AYB, FreeIAT). While the number of target categories is contingent on the research question, the test length depends on the respondents and the study environment. In an online experiment, individuals are more likely to drop out prematurely. A shorter test version might counter respondent fatigue, and should be preferred when addressing individuals with a lower attention span (such as children).

Data must be carefully prepared before it can be analysed. Specifically, error trials, extreme latencies, and practice blocks are potential sources of undesired variance. Extreme latencies can arise, for example, from the participants' lack of concentration. Possible treatments are trimming data and winsorizing data. In his seminal work on treating outliers within reaction time data, Ratcliff (1993) argues for cut-offs, which is a common practice when applying the improved algorithm (Greenwald et al. 2003). Researchers have the option to ignore, exclude, and recode errors, or they can analyse them separately. Error exclusion leads to information loss. Therefore, researchers usually recode error trials (+ 600 ms in the improved IAT algorithm).

Latency logarithmization addresses the issue of data skewness, which is generally observed in the reaction time data (Ratcliff 1993). Apart from calculating a D-Score (described in detail above), researchers can condense the reaction time data on an individual and a population level by calculating a Gaussian rank latency score (Sriram et al. 2006). This algorithm addresses the extreme latency problem because extreme latencies have less impact on a rank-based scoring. Robust statistics like the median value are also insensitive to outliers. The standard multivariate analysis methods (ANOVA and regression-based methods), as well as the quadruple pro-

cess model (Quad model) can disentangle four distinct task performance processes: association activation, the discriminability of correct response, success at overcoming bias, and the influence of guessing in the absence of other information (Conrey et al. 2005). Consequently, it is possible to distinguish the actual implicit attitude’s influence from other contaminating factors.

The IAT is the object of various criticisms (for an overview, see Teige-Mocigemba et al. 2010). Conceptually, Blanton et al. (2006) criticize the relative construct measurement. Fiedler et al. (2006) maintain that the diagnostic sign’s base rate (e. g., a high IAT score) exceeds the base rate of the attribute to be selected for diagnosis (e. g., definitely possessing a specific attitude toward an object). An environmentalist with a pro-environmental attitude might associate the otherwise cherished attitude object, environment, with negative elements, such as pollution and danger. Blanton et al. (2009) argue that the relationship between IAT scores and the behavioural outcome can be disappointingly small. Gawronski (forthcoming) maintains that researchers can expect large significant effects from the underlying dual-process models only in decision situations characterised by a low processing capacity.

Another critique states that there are alternative explanations for the observed latency differences, such as figure-ground asymmetry, salience asymmetries of the categories, task-set inertia, criterion shift, and stimulus-response compatibility. Furthermore, strategic effects, such as respondents’ recoding and cognitive skill, can confound the IAT scores (for an overview, see Teige-Mocigemba et al. 2010).

Finally, the IAT is criticized for causing method-specific variance, which manifests in correlations – even for non-

sense IATs (Mierke and Klauer 2003). Models of latency differences are inherently problematic, because fast responses in the compatible block are equated with slow responses in the incompatible block (Fiedler et al. 2006).

3.6. Case Illustration

The authors adapted an IAT with stimuli taken from Houben et al. (2009) whereby “Alcohol” and “Soft Drinks” functioned as the target categories, while “Active” and “Miserable” functioned as the attribute categories. Tab. 11 presents the procedural set-up and Tab. A2 the stimuli. The original stimulus material underwent a minor adaptation: Since the researchers found that the South African students are unfamiliar with the term “Long Drink,” this term was replaced with the term “Cocktail.” The IAT set-up consists of 180 trials for each respondent. The block sequence (with either the compatible or the incompatible tasks first) was varied to account for the possible order effects.

The authors condensed the IAT results into a D-Score according to the improved algorithm (Greenwald et al. 2003). They computed the D-Score such that a positive value indicates a stronger tendency to associate “Alcohol” with “Active.” The individual implicit associations were quite heterogeneous ranging from strongly negative to strongly positive (Tab. 12, right-hand column). On average, the participants were more inclined to associate “Alcohol” with “Miserable”, which the negative D-Score indicated.

A closer analysis of the latencies show that AIR and ADR are slower during the blocks that contain the “Alcohol” and “Miserable” combination (see Tab. 12 and Fig. 6). However, the difference is much more pro-

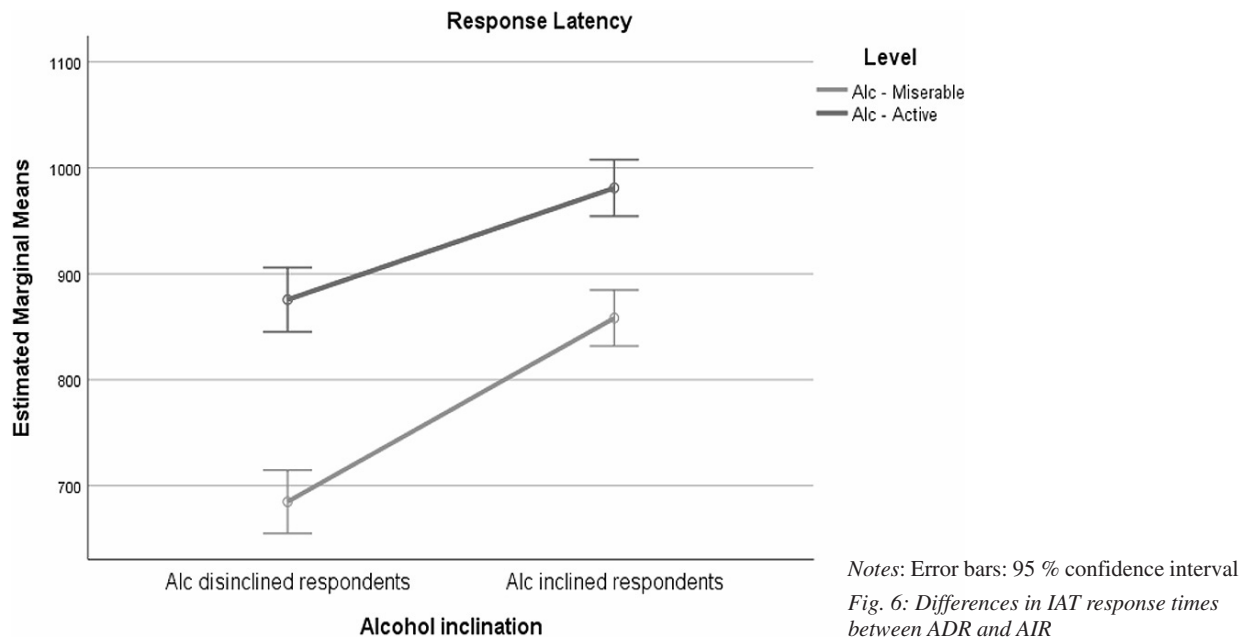
Tab. 11: IAT setup in the exemplary case application

Block	No. of trials	Category left	Category right
1	20	Alcohol	Soft Drinks
2	20	Active	Miserable
3a	20	Alcohol or Active	Soft Drinks or Miserable
3b	40	Alcohol or Active	Soft Drinks or Miserable
4	20	Soft Drinks	Alcohol
5a	20	Soft Drinks or Active	Alcohol or Miserable
5b	40	Soft Drinks or Active	Alcohol or Miserable

Tab. 12: Response latencies and D-Score

	ADR (n = 45)	AIR (n = 59)	Total (n = 104)
Latency (Alc – Miserable)	684.85 ± 19.23 (9.81)	858.22 ± 28.70 (14.64)	782.11 ± 18.31 (9.34)
Latency (Alc – Active)	875.58 ± 28.76 (14.68)	981.09 ± 31.15 (15.90)	935.00 ± 21.62 (11.03)
D-Score	- .46 ± 0.14 (.05)	- .24 ± .013 (.08)	- .34 ± 0.10 (.05)

Notes: Mean values ± 95 % confidence intervals, (standard error of the means) of alcohol-inclined (AIR) and alcohol-disinclined respondents (ADR)



nounced in respect of the ADR. Consequently, the D-Score is negative for both groups. The score is, however, more negative for the ADR than the AIR. The overall latency for the AIR is longer than for the ADR. The results demonstrate that the IAT is not only able to discriminate between groups (AIR and ADR), but can also predict a behavioural criterion (choice of alcohol). The authors performed a one-way ANOVA to demonstrate the difference between the AIR and the ADR with the analysis demonstrating significant differences between the two, $F(1,102) = 5.19, p = .02, \eta^2 = .484$.

To assess how implicit associations exert predictive power over decision-making, the authors joined the IAT with the DCE. The stimuli presented within the categories “Alcohol” and “Soft Drinks,” were used in exactly the same way as in the choice experiment. This approach justified using a simple regression model (see *Tab. 13*). The individuals’ percentage of alcoholic drinks chosen in the DCE functioned as a dependent variable (DV) and the D-Score as an independent variable (IV). The implicit arousal expectancies, which the D-Score measures (i.e., associating “Alcohol” with “Active”), predicted the percentage of alcoholic drinks in the choice experiment to a large extent. The IAT results explained a significant share of the variance. The results thus show that the IAT’s measured implicit associations are indeed linked to choices.

Model	B	S.E.	t	p	R ²
IAT D-Score	.22	.06	3.93	.000	.13

Notes: Dependent variable: Percentage of alcoholic drinks chosen.
S.E. = Standard Error

Tab. 13: Regression coefficients for IAT D-Score

4. Measuring Action Tendencies

Implicit cognition can cause a nonconscious predisposition of motoric behaviour toward an object (Chen and Bargh 1999, p. 216). Such action tendencies can be directed either at the stimulus (approach tendency) or away from it (avoidance tendency), and are deeply rooted in the human brain. Neuroscientific research revealed that approach and avoidance reactions are associated with distinct activation in the left and right prefrontal cortex (Roelofs et al. 2005). Reaction time (RT) paradigms, in which individuals immediately respond to the affective stimuli’s emotional valence by making arm movements that are either congruent or incongruent with their action tendencies, offer a relatively novel method of studying approach-avoidance tendencies (Roelofs et al. 2005). These methods do not explicitly ask the respondents to reveal their attitudes and are, thus, deemed more sufficient for examining human cognitive biases.

4.1. Research Paradigm

The common principle of AAT-based reaction time measurements is that the compatibility between the (motoric) response task and the stimuli’s valence affects the participants’ response speed (Heuer et al. 2007). Solarz (1960) demonstrated that there is an automatic inclination to approach desirable stimuli and avoid undesirable stimuli finding that people were faster to pull a lever toward them to indicate that a stimulus was positive and push a lever away from them to indicate that a stimulus was negative. Since then, researchers have applied these indirect measures of approach and avoidance tendencies successfully in different settings ranging from a traditional laboratory to a computer laboratory, and ultimately, to an online environment in which different devices, such as a physical lever, joystick, and mobile touch screen, are

used to study the cognitive aspects of psychological disorders.

The AAT was successfully applied to explain the approach and avoidance tendencies in different areas, such as fear (Heuer et al. 2007; Roelofs et al. 2010) and addiction (Wiers et al. 2010). For example, Rinck and Becker (2007) used the AAT to study the spider-fearful individuals' avoidance behaviour. Indirect measures were also used successfully to study the cognitive aspects of human action tendencies (Heuer et al. 2007). Wiers et al. (2010), as well as Woud et al. (2008), used the AAT for training purposes to overcome the implicit action tendencies of, for example, those with a phobia.

4.2. Applications in Marketing

Unlike the IAT, the AAT research has barely diffused into marketing literature. There are, however, reoccurring topics from related disciplines, most importantly from impulsive consumption research. Here, the AAT can be particularly insightful, since impulsive consumption often results from behavioural schemata that are performed without prior reflection when a stimulus is presented, for example, taking a cigarette when it is offered. In this respect, academics have studied responses towards food (Kakoschke et al. 2015; Lender et al. 2018; Verbeken et al. 2018), tobacco (Woud et al. 2016), and alcohol (Wiers et al. 2010). In a subsequent taste test, researchers also found that the AAT can predict unhealthy snack consumption (Kakoschke et al. 2015).

Moreover, researchers have used the AAT to change approach biases in human decision-making. In addition to the baseline AAT, consumers can be trained (AAT training) to avoid unhealthy items via more push trials that administer these stimuli and to approach healthy items via more pull trials administering these stimuli. Such AAT training and retraining have repeatedly proven effective in reducing approach bias and consumption in alcohol-related studies (Wiers et al. 2010). The training methods' effect on unhealthy snack consumption is yet ambiguous: While Becker et al. (2015), Dickson et al. (2016), as well as Verbeken et al. (2018), found that a modified approach bias did not affect consumption, a recent study by Schumacher et al. (2016) did trigger an effect.

While impulsive consumption accounts for the majority of marketing-related papers, certain researchers apply the AAT in the cognitive biases' realm, which is important to other consumer behaviour research fields. For example, Viswanathan et al. (2017) applied an AAT to obtain an approximate measure of the consumers' loss aversion. Researchers have also used the AAT to study attitudes toward social groups; however, to date these studies remain exclusively in the social psychology domain (Remedios et al. 2011).

Marketing studies are more widespread in the broader realm of approach-avoidance (A-A) behaviour. For ex-

ample, Summers and Hebert (2001) investigated A-A behaviour in the context of store illumination and Lee et al. (2011) did the same in the context of technology products. Here, the AAT method's unexploited application potential becomes evident.

4.3. Experimental Setup

4.3.1. Basic Scheme

Contrary to both the Stroop task's and IAT's measures, there is a less established and consistently shared AAT procedure (see *Tab. 14*). The AAT was historically measured in a conventional laboratory setting with an actual lever (Solarz 1960). Recent studies utilize computer-based instruments, for example, a joystick (Piqueras-Fiszman et al. 2014). In an AAT experiment, following a learning phase (practice level), the participants are asked to use the device to either pull an object (stimulus) toward or push it away from themselves. Depending on the experiment, the stimulus can consist of words, picture shapes, or only pictures. The number of stimuli in each category typically vary between six and 15 and can exceed 15. Overall, the entire experiment can run less than 80 or more than 200 trials.

Few researchers implement AAT in online settings (see e. g., Rinck and Becker 2007). Here, mouse movements replace the joystick. Alternatively, researchers can use a manikin task, during which the participants are required to move a manikin toward or away from objects displayed on a computer screen (Brignell et al. 2009). Very recently, smartphone-based solutions were implemented to present psychological experiments outside laboratory settings in a more ecologically valid environment.

Regardless of where the AAT is executed and which measurement instrument is used, these studies generally share the following set-up characteristics: The participants are confronted with an object (physically or on a screen) and are tasked with moving the object toward or away from themselves (or symbols of themselves). The responses are, thus, either compatible or incompatible with the behavioural tendencies to either approach or retract from the presented stimulus. Researchers generally run a few initial trials (9–12) so that the participants can learn the task, followed by a larger number of trials (around 80), during which a balanced set of compatible and incompatible situations are randomly presented.

4.3.2. Alternative Setups

Various visualizations are applied to enhance the AAT effect. Recent AAT modifications that rely directly on physical responses, and using them to measure approach and avoidance behaviours, are particularly promising. For example, the participants may be asked to move a figure, displayed on the screen, toward or away from a stimulus (Mogg et al. 2003). A comparison of the relative response speed functions as an implicit action tendency indicator. This approach successfully measured

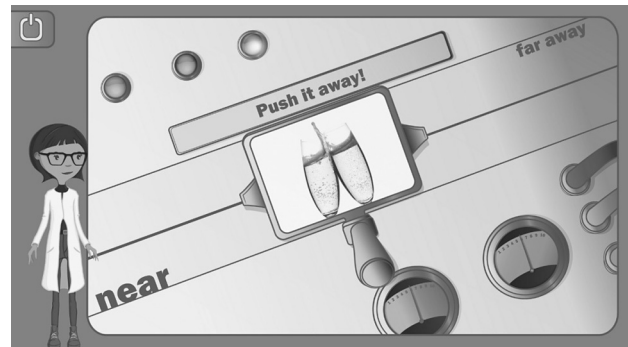
various substance dependencies (Field et al. 2006; Mogg et al. 2003), whereas its application in the marketing area is not (yet) known. The AAT experimental set-ups are very similar to the broader category of Stimulus-Response Compatibility (SRC) measures. These instruments measure the compatibility of a stimulus and the response features (Simon 1969). The higher the two conditions' compatibility, the shorter the expected response time. In this regard, the AAT is closely linked to the IAT.

The IAT procedure seeks to measure the implicit attitudes via their underlying automatic evaluation. Consequently, the procedure can be modified to measure automatic affect or action tendencies (Greenwald et al. 1998). For example, Ostafin and Palfai (2006), using a modified IAT, found that, alcohol-related stimuli are much more likely to trigger an approach behaviour than an avoidance behaviour in alcohol risky individuals. Schnabel et al. (2006) used an Implicit Association Procedure (IAP) as a hybrid between an IAT and an AAT. Similar to the IAT, the IAP aims to assess the automatic associations between concepts through a series of discrimination tasks. However, unlike the IAT, the IAP triggers automatic approach (pulling the joystick) and avoidance behaviour (pushing the joystick).

AAT experimental manipulations comprise congruent and incongruent tasks (Christiansen and Field 2013) with the incongruent tasks being more difficult to execute than the congruent ones. Consequently, when the task is modified from incongruent to congruent, the participants' attention level can change. This can be remedied by including compatible and incompatible stimuli in one block. Furthermore, the participants should not be instructed to make an explicit assessment of the stimuli (De Houwer 2003).

The reaction times regularly function as a dependent variable (Heuer et al. 2007; Seibt et al. 2007; Wiers et al. 2010). Alternatively, earlier cognitive researchers (see e. g., Bruyer and Brysbaert 2011) suggested that the inverse efficiency score (IES), which can be obtained from combining the percentage of errors and the reaction times, should be the measure of a dependent variable. Furthermore, Bartoszek and Winer (2015) used duration times, in addition to reaction times, in an AAT experiment of joystick movements.

To sum up, an approach-avoidance tendency can be observed by using different tasks in which the AAT seems to fit the best. Compared with other indirect (implicit) measures, such as memory association and attentional biases, the AAT has the unique advantage of directly measuring motoric action tendencies. The AAT also measures implicit cognitions more indirectly in the sense that the respondents are not required to act according to their implicit cognitions; instead, their attitudes are irrelevant to the task instructions (Rinck and Becker 2007). Thereby, the AAT can provide valuable insights that complement direct measures like the self-report measure (Heuer et al. 2007).



Notes: The participants were instructed to push the stimuli with a computer mouse (Web Experiment)

Fig. 7: Screenshot of an AAT trial using AYB

4.3.3. Online Implementation

The AAT is implemented as a two-stage process in the AYB research platform. First, the respondent explicitly evaluates a number of items by pressing a “thumbs up” or “thumbs down” icon. These selected items serve as a control and a benchmark for the experimental stimuli that the respondent does not rate explicitly.

The AAT experiment follows this step. Each target stimulus is presented in the centre of the screen. The respondents are asked to either pull the stimulus toward them or to push it away (tap or swipe to the left, or tap or swipe to the right). This push and pull task is conducted on a visual slider. A zooming effect supports the illusion of increasing or decreasing the distance toward the object. When pulled toward a respondent, the stimulus becomes larger and smaller when pushed away. Similar visualizations can be found in earlier studies (Chen and Bargh 1999; Kemps et al. 2013; Seibt et al. 2007). The participants are asked to respond as quickly and as accurately as possible. Immediate feedback via an avatar and a test score interpretation at the end of the experiment is also provided. Fig. 7 above shows an AAT trial using an AYB platform.

4.4. Data Analysis Steps

A mixed ANOVA is used to examine the differences between the median response time (RT) and Error rates of the push and pull conditions. The RTs of the correct responses are used as the dependent variable; median RTs are determined for each participant and for each of the picture type and response direction combinations (Heuer et al. 2007).

$$\text{Bias (AAT) scores} = \text{RTs median (push conditions)} - \text{RTs median (pull conditions)} \quad (4)$$

Note: Relative strength of approach-avoidance tendency: Negative values (scores) = Higher negative reactions (stronger avoidance)

Response latency is commonly used as a dependent variable, partly complemented by an analysis of error rates (Heuer et al. 2007; Rinck and Becker 2007; Seibt et al. 2007; Wiers et al. 2010). Seibt et al. (2007) also interpret

the absolute speed of approach-avoidance responses as an indicator of the respondents' need state and impulsivity. Moreover, AAT effect scores can be computed by subtracting each participant's median RT in the pull condition from the median RT in the corresponding push condition. A negative AAT score indicates an avoidance bias, while a positive AAT score indicates an approach bias (Heuer et al. 2007).

Similar to other research on visual probe tasks (Field et al. 2006), RTs are excluded if they are less than a lower threshold (200 ms – 300 ms), since these responses are considered anticipated responses. They are also eliminated if they are greater than an upper threshold (2000 ms) and if they are more than 3 SDs above the respondents' mean latency. Since the latencies of the incorrect responses are difficult to interpret, only the correct responses are used in this analysis (Heuer et al. 2007).

To assess the provided answers' reliability, split-half reliability scores can be calculated using a procedure similar to the procedure outlined in the IAT section. Researchers should be careful if they observe a sign reversal when comparing the first half block's AAT score with the second half set of reactions' corresponding score.

4.5. Methodological summary and limitations

Tab. 14 provides an overview of the AAT methodological procedures. The AAT can vary in a number of ways, such as the configuration, study set-up, data treatment, and data analysis procedures.

While the task instruction plays a crucial role in evoking respondents' reactions, there is no consensus in the literature on different task instructions' effectiveness. Studies ask the participants to react to content-irrelevant stimuli, such as stimuli frames, shapes, and colours. Hereby, a mixture of verbal and nonverbal instructions might be a favourable solution. Finally, previous studies (Greenwald et al. 1998, p. 1467) suggest that “the values in these tails of the latency distribution are problematic not only because they lack theoretical interest but also because they distort means and inflate variances.” Consequently, data-cleaning, as well as the dependent measures' aggregation and disaggregation, should be done carefully.

The indirect tasks' limitation is that they mainly tap into emotional information processing's semantic aspects (Rinck and Becker 2007). The AAT has specific methodological limitations, mainly “its use of behavioural ten-

Configuration	Environment	Traditional laboratory		Computer laboratory*		Online
	Input device	Physical Lever		Joystick / Mouse*		Mobile/ Touchscreen*
	Software	Basic freeware	Nonprofit software	Commercial software	Apps*	
Study setup	Stimuli categories	Congruent + control		Congruent + incongruent + control		Congruent + incongruent + control + neutral*
	Stimuli types	Stimulus shape		Pictures		Pictures overlaid with words*
	Stimuli selection	IPS		Pure Stimuli		Embedded stimuli
	Number of Stimuli/ category	Low (< 6 stimuli)		Medium (< 15 stimuli)		High (≥ 15 stimuli)*
	Test length	Brief (less than 80 trials)*		Regular (104 trials)		Long (> 200 trials)*
Data treatment	Practice trials	Exclude first block of trials (9-12 trials)		Exclude first trial(s) in each block		No exclusion of any trials
	Error treatment	No treatment			Exclude errors	
	Extreme observations	No treatment		Anticipated response		Errors
	Differentiation lower & upper limit	Below 150 ms and higher than 1450 ms		Below 250 ms and higher than 2000 ms		Below 300 ms and higher than 3000 ms
Data analysis (standard software)	Dependent variable	Reaction time (Response latency)	Duration time (AAT-DT)*	Error rates		The inverse efficiency score (IES)*
	Aggregation level for DV	Population mean values		Individual mean values		AAT-Score*
	Analysis method	ANOVA		Repeated ANOVA		Mixed Models
	Reliability checks	Descriptive statistics/Internal consistency		Split-half reliability scores		Criterion (predictive) validity

Tab. 14: Methodological conventions of AAT: Traditional default in grey, AYB study example in italics, *) available options within core AYB program

dencies as an indicator of attitudes” (Klein et al. 2011; Rinck and Becker 2007). The approach and avoidance tendencies could reach beyond the pull and push conditions. Neither of the two arm movements might be an “ideal operationalization” (Kozlik et al. 2015). The standard assumption is that a static participant moves a stimulus and, therefore, arm extensions signal a push condition. However, the stimuli might be static and the participant might move, making an arm extension mean the opposite, implying a movement toward the object.

Furthermore, the task “does not distinguish between different types of attitude” (Rinck and Becker 2007, p. 118). For example, the negative attitude (push) could be related to fear and disgust of the stimuli. The AAT, thus, not necessarily reflects stimulus valence, because an overall positive stimulus can sometimes activate an avoidance reaction (Klein et al. 2011).

4.6. Case Illustration

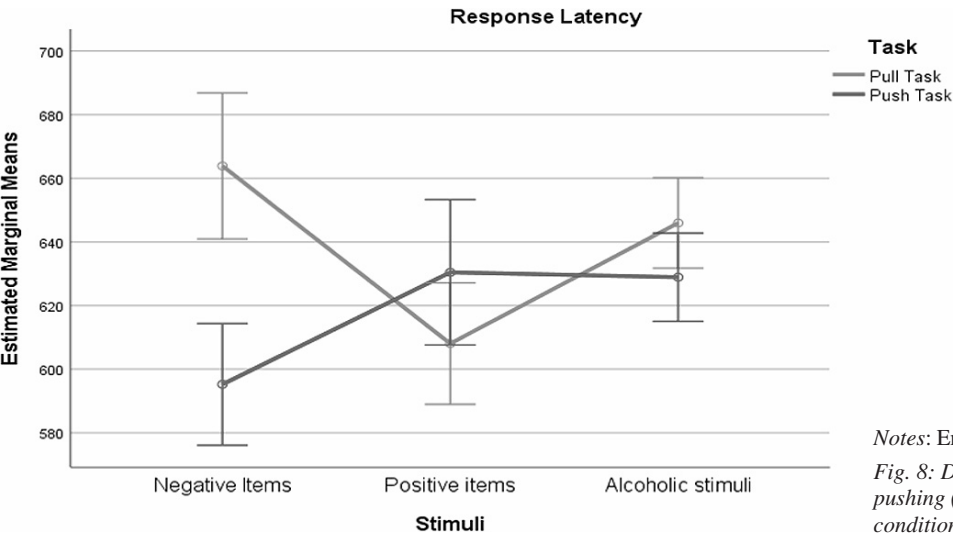
The AAT was implemented with images of the stimuli used in the IAT and taken from Houben et al. (2009). The participants received short written instructions with illustrations on what was expected from them during the entire task. In the first part of the experiment (optional in the AYB platform), a set of 12 food-related control stimuli consisting of equal amounts of fast food and healthy food, as well as fresh food and rotten food, was shown to the participants. Each participant was asked to select three “positive” and three “negative” control stimuli. This personal selection procedure ensured that the control stimuli would relate predictably to a person-specific approach-avoidance behaviour.

In the AAT experiments’ first level, only the control stimuli were used to instruct the participants on how the test works. The AAT task itself started with ten practice trials in which the participants learned to push or pull the joystick. The instruction was always consistently held in the practice phase, allowing the respondents to internalize the specific stimulus movement’s meaning.

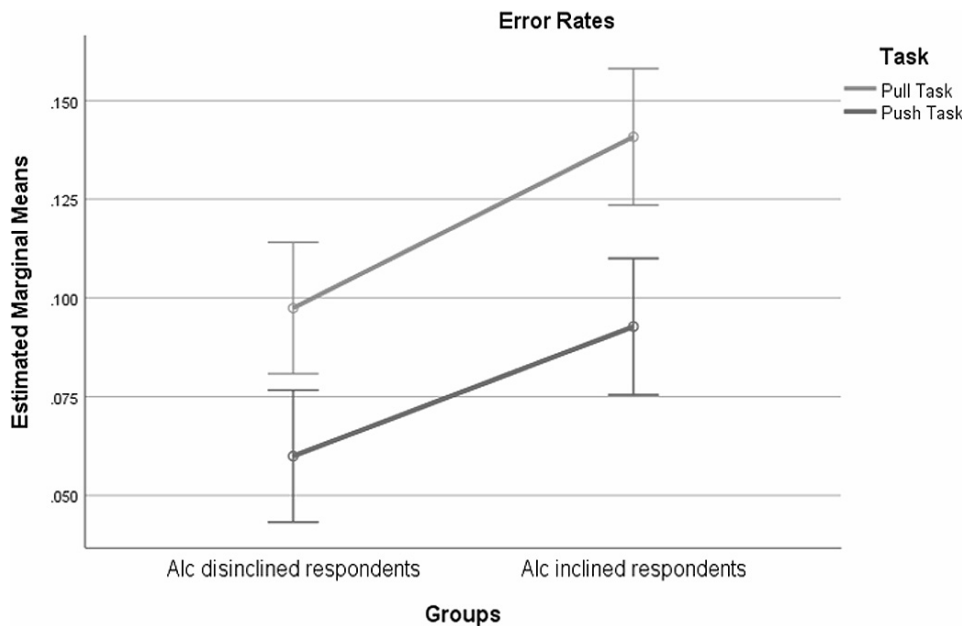
The test phase consisted of 104 trials in which an equal number of pictures of alcoholic drinks and non-alcoholic drinks were displayed in either a push or pull condition. They were randomly interrupted by control stimuli in one of the two task conditions. Following Heuer et al. (2007), the order of the experimental blocks was counter-balanced across the participants and each participant completed 104 trials (ten practice trials). The independent variables (response latency and error rates) were analysed in a multivariate ANOVA design. The task (push vs. pull), stimulus (positive vs. negative), and stimuli categories were coded as fixed factors, which yielded a 2 x 2 x 2 factorial design.

In accordance with the procedure commonly applied in prior studies, the response latency was used as the primary dependent variable (Heuer et al. 2007; Seibt et al. 2007; Wiers et al. 2010). Errors (9.5 %) and RTs smaller than 200 ms (1.2 %) or larger than 2000 ms were excluded. The lower threshold eliminates the trials that were not meaningfully processed (Greenwald et al. 1998; Posner 1980). Various thresholds were tested as a robustness check, but these tests had no impact on the statistical effects. The control items were analysed to verify that the AAT test procedure had functioned adequately.

The results of the reaction time analyses reveal that the task exerts a significant main effect and that there is a significant interaction effect between the stimuli and the task. The ANOVA analysis of the response latency reveals significant differences between the two groups of participants with $F(1,7489) = 92.72, p < .001, \eta^2 = .012$. Furthermore, the results imply that the pull and push conditions (task) differ significantly between the groups, $F(1,7489) = 7.087, p = .008, \eta^2 = .001$. The interaction effect between the stimuli category and the task also discloses a significant finding, $F(2,7489) = 9.046, p < .001, \eta^2 = .002$. Fig. 8 illustrates that the participants responded faster when pushing negative control stimuli than when pulling them. The effect for the positive items is, as expected, reversed, while the differences between the push and pull conditions are not significant. Finally,



Notes: Error bars: 95 % confidence interval
Fig. 8: Differences in AAT RT between pushing (avoid) and pulling (approach) conditions



Notes: Error bars: 95 % confidence interval

Fig. 9: Differences in AAT error rates between pushing (avoid) and pulling (approach) alcoholic stimuli

the push and pull latencies do not differ in respect of the displayed alcohol stimuli, suggesting that these stimuli were on average, positioned as neutral.

A comparison between the two subgroups of alcohol inclined and disinclined respondents reveals pronounced findings for error rates as dependent variable. Fig. 9 illustrates the error rate differences between pushing (avoid) and pulling (approach) the alcoholic stimuli. In both groups the incongruent tasks were more prone to errors than the congruent tasks, because the respondents pulling negative items to themselves (incongruent task) induced a higher error rate than when pulling positive items to themselves (congruent task). Conversely, pushing a positive item away (incongruent task) provoked a higher error ratio than pushing negative items away.

An ANOVA with error rate as the dependent variable, and task, and stimulus, as independent factors discloses further evidence of the effects. The findings show significant differences between the two groups of participants with $F(1,8255) = 25.32, p < .001, \eta^2 = .003$. The results also imply significant differences between the groups' pull and push condition (task) with $F(1,8255) = 7.706, p = .006, \eta^2 = .001$.

Our case application's exemplary findings provide initial evidence that error rates can function as a dependent variable in reaction time paradigms. Evidence was obtained in respect of alcohol-inclined respondents, thereby supporting the expected predisposition. Our findings indicate that this approach bias may well be prevalent in regular consumption-related settings.

Finally, regression analyses (Tab. 15) were performed to clarify the direct effects on the consumers' real choices and to assess the AAT scores' predictive validity. Unfortunately, a first model ((1) in Tab. 15), which was based on the AAT score of all the alcohol-related stimuli, ex-

Model	B	S.E.	t	p	R ²
1. AAT score for all stimuli	.006	.00	.65	.94	.00
2. AAT score for hard liquor	-.23	.00	-2.42	.017	.06

Notes: Dependent variable: Percentage of alcoholic drinks chosen. S.E. = Standard Error

Tab. 15: Regression coefficient for AAT scores

hibits no relationship between the AAT and the DCE choices. Again, preference heterogeneity might have caused this, as the respondents might not approach or avoid all the different alcohol stimuli to the same extent. We, therefore, restricted the analysis to hard liquor stimuli as a proxy for more severe alcohol drinking habits. This restriction to severe alcohol stimuli seems justified, since previous research focused on addiction issues. The second model ((2) in Tab. 15) builds on a revised AAT score calculated solely on the basis of the respondents' responses to hard liquor stimuli. This narrower specification of alcohol-related approach tendencies led to the expected finding: The AAT score's significant negative regression coefficient implies that the percentage of alcoholic items chosen in the DCE increases if the respondents pull alcohol stimuli faster toward themselves than pushing them away.

5. General Discussion

5.1. Overall reflections

Conventional measures of consumers' insight are implemented via interviews, questionnaires, and surveys. Measuring the customers' implicit cognitions sufficiently via these approaches is, however, difficult. These approaches are deemed insufficient, owing to self-presentation strategies, response task characteristics, distortions

based on social desirability, and their tendency to address decision outcomes without focusing on the processes that precede a decision (Heuer et al. 2007). Indirect cognition measures overcome the explicit survey instruments’ limitations. These measures enable an in-depth process perspective on decision-making and can, therefore, reveal the underlying causes of consumer actions. Actions and decisions are automatically activated without the performer being aware of their causation (Greenwald et al. 1998). In this regard, the Stroop test and the AAT can be considered more indirect tests than the IAT, since the latter requires the participants to categorize the stimuli content in order to respond.

The three methods described in this paper evaluate different cognitive processes, although these methods are all based on the measurement of response latencies. The Emotional Stroop measures cognitive distraction based on the displayed stimuli’s emotional valence. The higher the stimuli’s valence, the longer the reaction time. The IAT assesses the relative strength of associations between cognitive elements. The higher the associative strength, the faster the compatible and the slower the incompatible categorization tasks. The AAT adds response effects at the motoric level. The respondents pull the approach-inducing stimuli faster toward themselves and push them away more slowly, than they do when confronted with avoidance-inducing stimuli.

5.2. Marketing applications

The authors argue throughout this paper that implicit cognition measures have a large potential regarding marketing applications. Each implicit measure has unique characteristics that render it, more or less, suited for specific applications. *Tab. 16* provides an overview to support researchers in their decision on which measure to use. The application fields are broadly categorized into

visceral and biological processes, attention-based processes, and higher cognitive processes.

The visceral and biological influences include drive states, such as hunger, thirst, sexual desire, emotions, physical pain, and drug cravings (Loewenstein 1996). These triggers have an immediate impact on an individual’s hedonic experience and alter the perceptions of desirability in respect of products and actions (Loewenstein 1996). As such, an analysis of visceral biological influences might be of interest with regard to impulsive consumption and the effects of, for example, nudity in advertising, food marketing, and sensory marketing.

Approach-avoidance behaviour is rooted in one of the most fundamental appetitive/aversive biological reactions – prevalent even in simple organisms (Elliot 1999). Therefore, the AAT is arguably the most suitable method for measuring bodily reactions towards stimuli which requires little or no deliberation. Since substance abuse can alter neurological circuits such that drug-related stimuli become more salient (Robinson and Berridge 2001), implicit attention measures, such as the Emotional Stroop task, are important for addiction research. This reasoning applies to a lesser degree to nutritional stimuli. The IAT is based on the semantic understanding and the associative activation of words which is why we deem the IAT less applicable in this regard.

Attention constitutes a selective processing that determines to which stimuli or actions a limited processing capacity is devoted. As such, attentional biases might apply to settings in which consumers can or cannot pay attention to marketing stimuli, for example, in a retail or advertising context. Bottom-up and top-down processes determine attention, with the former directing attention unintentionally towards an object based on the stimuli’s salience and biological importance (Robinson and Berridge 2001). However, top-down processes, especially

Exemplary marketing applications		Approach-avoidance / AAT	Implicit attention / Stroop	Implicit association / IAT
Visceral and biological processes	Impulsive consumption Addiction Nudity in advertising Sensory Marketing Food marketing ...	Yes	Yes, especially addiction and food marketing	Yes, as long as associative processes are involved
Attention focused processes	Ad exposure Retail positioning Sports sponsorship Brand logo Online banners ...	No, no action required	Yes	Can measure only outcomes of attention based processes, not process itself
Higher cognitive processes	Brand attitude Preference formation Consumption goals Brand personality theories Consumer decision-making ...	AAT is unidimensional, HCP are multidimensional	Some HCP, e. g. goals, guide attention However, implicit attention is not sufficient for HCP	Yes

Tab. 16: Marketing application potential for implicit cognition measures

goals, drive most of the attentive processes (Dijksterhuis and Aarts 2010). For example, drinks attract more attention on a supermarket shelf if one is thirsty. The outcome of an implicit attention measure, such as the Emotional Stroop task, captures implicit attention bias and can be indicative, but not diagnostic, of higher cognitive processes, such as goals.

Finally, higher cognitive processes rely more on the cortical brain areas and require a larger amount of deliberation (Dijksterhuis and Aarts 2010). These higher cognitive processes come into play whenever consumers engage in volitional behaviour or during deliberate mental operations, such as deciding on goals, attitudes, preferences, and judgements. Goals (in their most fundamental form) can be drawn to approach-avoidance behaviour (Elliot 1999). The same cannot be said of multidimensional constructs, such as attitudes. In this instance, the IAT is most applicable method to capture such higher cognitive processes.

5.3. Practical suggestions for marketing researchers

The indirect measurement approaches require an especially careful implementation of the experimental set-up, because noise, as well as systematic measurement errors, might dilute the findings. This dilution explains the often low correlations between repetitions, as well as the low correlations between different implicit cognition measures (Fazio and Olson 2003). Noise is inherent in implicit cognition tasks; this is understandable, given that the respondents are not asked to answer challenging questions carefully. Systematic errors might occur, due to flaws in the experimental set-up such as badly fitting stimuli. Whenever possible, researchers should, therefore refer to established experimental set-ups and well-tried stimuli. We recommend in-depth manipulation checks in the target population to ensure that one truly measures the phenomenon of interest. Our Stroop application serves to a certain extent as a negative example. We applied established items in their original language, but were confronted with an unexpected ambiguity of certain nouns in the target population.

Special care should be given to the treatment of outliers. Researchers familiar with analyzing closed-ended questions and Likert-type scales need to investigate the retrieved data's distributional patterns carefully before applying mean comparisons of the effect sizes. While there are well-reported cut-points for acceptable lower and upper latency values of the various test procedures, these values should be taken as rules of thumb and only be used as a rough guideline. Sensitivity analyses for the alternative handling of outliers should help verify the findings' robustness. Obviously, large-scale research is necessary to overcome the potential pitfalls of the outliers' strong effects.

Implicit cognition measures' high context dependency needs to be asserted on a more general level. The quest

for ubiquitous test instruments seems to be unrealistic. Instead, researchers should follow the C-OAR-SE framework: construct definition, object classification, attribute classification, rater identification, scale formation, and enumeration (Rossiter 2002). Our application case exemplifies this need for careful upfront considerations of the study's scope: If we were to repeat the study, we would focus on a single subcategory like hard liquor.

Social scientists conventionally apply implicit cognition measures to compare cognitions across extreme settings and groups; addiction behavior is a good example. These settings have to be carefully adapted for marketing research purposes. Consumption-related issues are generally not crucially important to individuals; consequently, one cannot expect the implicit cognitions under investigation – whether awareness, associations, or even action tendencies – to cause extreme reactions of study participants. Instead, consumer behavior researchers will quite likely have to deal with comparably small effect sizes, similar to those reported in the application case. This requires larger scale sample sizes than those used for research on generic issues of social psychology. Empirical research on implicit cognition will, consequently, have to leave its traditional laboratory environment to gain the reliable and representative insights needed for marketing practice. Conducting implicit cognition research online constitutes a promising opportunity, since it can combine the advantages of retrieving and handling large-scale data with the ease of answering for the participants.

5.4. Potential for Social Marketing

Understanding consumer impulsivity and the implicit cognitions' role in decision-making is important not only for marketers, but also for consumer protection. Behavior based on implicit cognitions frequently makes people behave in a way that economic theory finds difficult to predict and understand. Thaler and Sunstein (2009) therefore incorporated psychologically realistic assumptions into economic decision-making analyses and proposed the nudge theory to improve decisions about health, wealth, and happiness. On a more extreme level, psychopathology refers to a combination of disorder-specific impulsive processes and weak general abilities to control these impulses via reflective longer-term considerations (Wiers et al. 2013). Further implicit cognition research may help consumers quit addictive and impulsive tendencies.

Future applications of implicit cognition measures need not remain restricted for diagnostic purposes, but should be applied in novel settings to invoke attitude change for the better. Wiers et al. (2013) distinguish between two types of training interventions. Cognitive bias modification (CBM) aims at changing maladaptive cognitive motivational biases, whereas working memory training aims at training domain-general cognitive control. Combinations of motivational intervention and cognitive training are especially promising (Wiers et al. 2015). However, little is known about successful training (nudging),

which underpins the potential of (neuro)cognitive dual-process models for sustainable consumption in marketing and consumer research.

Initial research indicates that implicit cognitions can be modified by using the Stroop tasks, IATs, or AATs, and that this can influence subsequent consumption. For example, approach bias modification seems effective to reduce approach bias and unhealthy consumption behavior (Becker et al. 2015). Combinations of motivational intervention and cognitive training are the most promising (Wiers et al. 2015), but to date the evidence does not provide a solid basis for applying implicit cognition trainings (Dickson et al. 2016) and more research is required.

Appendix

	1st	2nd	3rd	4th	5th
Route I	Stroop	Posner	IAT	AAT	DCE
Route II	Posner	AAT	Stroop	IAT	DCE
Route III	AAT	IAT	Posner	Stroop	DCE
Route IV	IAT	Stroop	AAT	Posner	DCE

Tab. A1: Order of implicit cognition tests

References

Ackermann, C.-L., Teichert, T., & Truong, Y. (2018). ‘So, What Is It? And Do I Like It?’ New Product Categorisation and the Formation of Consumer Implicit Attitude. *Journal of Marketing Management*, 34, 796–818.

Algom, D., Chajut, E., & Lev, S. (2004). A Rational Look at the Emotional Stroop Phenomenon: A Generic Slowdown, Not a Stroop Effect. *Journal of Experimental Psychology: General*, 133, 323–338.

Anderson, J. R. (1983). A Spreading Activation Theory of Memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261–295.

Banaji, M. R. (2013). The Implicit Association Test at Age 7: A Methodological and Conceptual Review. In Bargh, J. (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*, Ney York: Psychology Press, 265–292.

Bargh, J. A. (2002). Losing Consciousness: Automatic Influences on Consumer Judgment, Behavior, and Motivation. *Journal of Consumer Research*, 29, 280–285.

Bartoszek, G., & Winer, E. S. (2015). Spider-Fearful Individuals Hesitantly Approach Threat, Whereas Depressed Individuals Do Not Persistently Approach Reward. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 1–7.

Becker, D., Jostmann, N. B., Wiers, R. W., & Holland, R. W. (2015). Approach Avoidance Training in the Eating Domain: Testing the Effectiveness across Three Single Session Studies. *Appetite*, 85, 58–65.

Bellis, M. A., Phillips-Howard, P. A., Hughes, K., Hughes, S., Cook, P. A., Morleo, M., Hannon, K., Smallthwaite, L., & Jones, L. (2009). Teenage Drinking, Alcohol Availability and

Notes

- [1] The search phrases combined the measurement methods with their application fields. We differentiated between two application fields: (a) an application field within the broader research field of behavioral decision-making (“decision-making” OR “implicit cognition*”) and (b) an application field within marketing-related applications (“marketing*” OR “customer*” OR “consumer*” OR “advertise*” OR “brand*”). Method-wise, we searched for “Stroop*” and “implicit association test.” We applied a narrow and a wider search phrase to identify research on approach and avoidance tendencies, because academics have not yet established a standard procedure and wording.
- [2] Since the original study (Stormark et al. 1997) was conducted in Norway, two UK-specific measurement words were excluded from the set of control words: “gram” and “pound.” In Hindu (there are many Indian students in SA) “gram” also means village and in SA “pound” is predominantly considered a currency; “pound” may, therefore, confuse participants or rouse their attention.

Object Category		Association Category	
Alcohol	Soft Drinks	Active	Miserable
Beer	Fanta	Cheerful	Cheerless
Wine	Coke	Energetic	Horrible
Whisky	Sprite	Lively	Nauseous
Cocktail	Pepsi	Excited	Unhappy
Vodka	Ice-Tea	Chatty	Listless

Tab. A2: IAT stimuli in case application

Pricing: a Cross-Sectional Study of Risk and Protective Factors for Alcohol-Related Harms in School Children. *BMC Public Health*, 9, p. 380.

Berry, T. R. (2006). Who’s Even Interested in the Exercise Message? Attentional Bias for Exercise and Sedentary-Lifestyle Related Words. *Journal of Sport and Exercise Psychology*, 28, 4–17.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for Criterion Prediction. *Journal of Experimental Social Psychology*, 42, 192–212.

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT. *Journal of Applied Psychology*, 94, 567–82.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the Perfect Measure of Implicit Self-Esteem: The Blind Men and the Elephant Revisited? *Journal of Personality and Social Psychology*, 79, 631–643.

Brandimonte, M., Bruno, N. N., & Collina, S. (2006). Cognition. In P. Pawlik & G. d’Ydewalle (Eds.), *Psychological Concepts: An International Historical Perspective*. Hove UK: Psychology Press, 11–26.

Braun-LaTour, K. A., Puccinelli, N. M., & Mast, F. W. (2007). Mood, Information Congruency, and Overload. *Journal of Business Research*, 60, 1109–1116.

Brignell, C., Griffiths, T., Bradley, B. P., & Mogg, K. (2009). Attentional and Approach Biases for Pictorial Food Cues. Influence of External Eating. *Appetite*, 52(2), 299–306.

Brough, A. R., Wilkie, J. E. B., Ma, J., Isaac, M. S., & Gal, D. (2016). Is Eco-Friendly Unmanly? The Green-Feminine Stereotype and Its Effect on Sustainable Consumption. *Journal of Consumer Research*, 43, 567–582.

- Brown, H. R., Zeidman, P., Smittenaar, P., Adams, R. A., McNab, F., Rutledge, R. B., & Dolan, R. J. (2014). Crowdsourcing for Cognitive Science – the Utility of Smartphones. *PloS one*, 9(7): e100662, 1–9.
- Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a Valid and Valuable Measure of Implicit Consumer Social Cognition? *Journal of Consumer Psychology*, 14(4), 385–404.
- Bruyer, R., & Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a Better Dependent Variable than the Mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychologica Belgica*, 51, 5–13.
- Camerer, C., Loewenstein, G. F., & Rabin, M. (Eds.). (2004). *Advances in Behavioral Economics* (The Roundtable Series in Behavioral Economics). New York, Princeton, Oxford: Russell Sage Foundation; Princeton University Press.
- Chen, M., & Bargh, J. A. (1999). Consequences of Automatic Evaluation: Immediate Behavioral Predispositions to Approach or Avoid the Stimulus. *Personality and Social Psychology Bulletin*, 25(2), 215–224.
- Christiansen, P., & Field, M. (2013). Implicit Cognition. In J. MacKillop & H. de Wit (Eds.), *The Wiley-Blackwell Handbook of Addiction Psychopharmacology*. Chichester, West Sussex, Malden, Mass: Wiley-Blackwell, 489–514.
- Cian, L., Krishna, A., & Schwarz, N. (2015). Positioning Rationality and Emotion: Rationality Is up and Emotion Is Down. *Journal of Consumer Research*, 42, 632–651.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating Multiple Processes in Implicit Social Cognition: the Quad Model of Implicit Task Performance. *Journal of Personality and Social Psychology*, 89, 469–487.
- Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The Addiction-Stroop Test: Theoretical Considerations and Procedural Recommendations. *Psychological Bulletin*, 132, 443–476.
- De Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology*, 50, 77–85.
- De Houwer, J., & Bruycker, E. de. (2007). The Implicit Association Test Outperforms the Extrinsic Affective Simon Task as an Implicit Measure of Inter-Individual Differences in Attitudes. *British Journal of Social Psychology*, 46(2), 401–421.
- De Houwer, J., & Moors, A. (2010). Implicit Measures: Similarities and Differences. *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*, 176–193.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit Measures: A Normative Analysis and Review. *Psychological Bulletin*, 135, 347–368.
- Dickson, H., Kavanagh, D. J., & MacLeod, C. (2016). The Pulling Power of Chocolate: Effects of Approach-Avoidance Training on Approach Bias and Consumption. *Appetite*, 99, 46–51.
- Dijksterhuis, A., & Aarts, H. (2010). Goals, Attention, and (Un)consciousness. *Annual Review of Psychology*, 61, 467–490.
- Dimofte, C. V. (2010). Implicit Measures of Consumer Cognition: A Review. *Psychology and Marketing*, 27, 921–937.
- Dong, R., & Gleim, M. R. (2018). High or Low: The Impact of Brand Logo Location on Consumers Product Perceptions. *Food Quality and Preference*, 69, 28–35.
- Elliot, A. J. (1999). Approach and Avoidance Motivation and Achievement Goals. *Educational Psychologist*, 34, 169–189.
- Fadardi, J. S., & Cox, W. M. (2009). Reversing the Sequence: Reducing Alcohol Consumption by Overcoming Alcohol Attentional Bias. *Drug and Alcohol Dependence*, 101, 137–145.
- Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, 54, 297–327.
- Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved Problems with the “I”, the “A”, and the “T”: A Logical and Psychometric Critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, 17, 74–147.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in Automatic Activation as an Unobtrusive Measure of Racial Attitudes: A Bona Fide Pipeline? *Journal of Personality and Social Psychology*, 69, 1013–1027.
- Field, M., & Cox, W. M. (2008). Attentional Bias in Addictive Behaviors: a Review of Its Development, Causes, and Consequences. *Drug and Alcohol Dependence*, 97, 1–20.
- Field, M., Christiansen, P., Cole, J., & Goudie, A. (2007). Delay Discounting and the Alcohol Stroop in Heavy Drinking Adolescents. *Addiction*, 102, 579–586.
- Field, M., Eastwood, B., Bradley, B. P., & Mogg, K. (2006). Selective Processing of Cannabis Cues in Regular Cannabis Users. *Drug and Alcohol Dependence*, 85, 75–82.
- Florack, A., Friese, M., & Scarabis, M. (2010). Regulatory Focus and Reliance on Implicit Preferences in Consumption Contexts. *Journal of Consumer Psychology*, 20, 193–204.
- Forehand, M. R., & Perkins, A. (2005). Implicit Assimilation and Explicit Contrast: A Set/Reset Model of Response to Celebrity Voice-Overs. *Journal of Consumer Research*, 32, 435–441.
- Freeman, J. B. (2014). Abrupt Category Shifts during Real-Time Person Perception. *Psychonomic Bulletin & Review*, 21(1), 85–92.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for Studying Real-Time Mental Processing Using a Computer Mouse-Tracking Method. *Behavior Research Methods*, 42(1), 226–241.
- Friese, M., Wänke, M., & Plessner, H. (2006). Implicit Consumer Preferences and Their Influence on Product Choice. *Psychology and Marketing*, 23, 727–740.
- Frings, C., Englert, J., Wentura, D., & Bermeitinger, C. (2010). Decomposing the Emotional Stroop Effect. *The Quarterly Journal of Experimental Psychology*, 63, 42–49.
- Fukawa, N., & Niedrich, R. W. (2015). A Fluency Heuristic Account of Supraliminal Prime Effects on Product Preference. *Psychology & Marketing*, 32, 1061–1078.
- Gawronski, B. (forthcoming). Six Lessons for a Cogent Science of Implicit Bias and its Criticism. *Perspectives on Psychological Science*.
- Gawronski, B., & Bannon, S. M. (2019). Attitudes and the Implicit-Explicit Dualism. In D. Albarracín & B. T. Johnson (Eds.), *The Handbook of Attitudes: 2nd edition*. New York, NY: Taylor & Francis, 158–196.
- Graf, P., & Schacter, D. L. (1985). Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 501–518.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes Self-Esteem, and Stereotypes. *Psychological Review*, 102(1), 4–27.
- Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie*, 48(2), 85–93.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition: the Implicit Association Test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity. *Journal of Personality and Social Psychology*, 97, 17–41.
- Gregg, A. P., & Klymowsky, J. (2013). The Implicit Association Test in Market Research: Potentials and Pitfalls. *Psychology & Marketing*, 30, 588–601.
- Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced Mouse-Tracking Analytic Techniques for Enhancing Psychological Science. *Group Processes & Intergroup Relations*, 18, 384–401.

- Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of Emotional Facial Expressions in Social Anxiety: The Approach-Avoidance Task. *Behaviour Research and Therapy*, 45, 2990–3001.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation between the Implicit Association Test and Explicit Self-Report Measures. *Personality & Social Psychology Bulletin*, 31, 1369–1385.
- Hornung, J., Kogler, L., Wolpert, S., Freiherr, J., & Derntl, B. (2017). The Human Body Odor Compound Androstadienone Leads to Anger-Dependent Effects in an Emotional Stroop but not Dot-Probe Task Using Human Faces. *PloS one*, 12(4): e0175055, 1–19.
- Houben, K., Rothermund, K., & Wiers, R. W. (2009). Predicting Alcohol Use with a Recoding-Free Variant of the Implicit Association Test. *Addictive Behaviors*, 34, 487–489.
- Kahneman, D. (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist*, 58, 697–720.
- Kakoschke, N., Kemps, E., & Tiggemann, M. (2015). Combined Effects of Cognitive Bias for Food Cues and Poor Inhibitory Control on Unhealthy Food Intake. *Appetite*, 87, 358–364.
- Kalanthroff, E., Henik, A., Derakshan, N., & Usher, M. (2016). Anxiety, Emotional Distraction, and Attentional Control in the Stroop Task. *Emotion (Washington, D.C.)*, 16, 293–300.
- Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a Measure of Implicit Social Cognition. *Journal of Personality and Social Psychology*, 91, 16–32.
- Keller, K. L., Kuilema, L. G., Lee, N., Yoon, J., Mascaro, B., Combes, A.-L., Deutsch, B., Sorte, K., & Halford, J. C. G. (2012). The Impact of Food Branding on Children's Eating Behavior and Obesity. *Physiology & Behavior*, 106, 379–386.
- Kemps, E., Tiggemann, M., Martin, R., & Elliott, M. (2013). Implicit Approach-Avoidance Associations for Craved Food Cues. *Journal of Experimental Psychology: Applied*, 19(1), 30–38.
- King, M. F., & Bruner, G. C. (2000). Social Desirability Bias: A Neglected Aspect of Validity Testing. *Psychology and Marketing*, 17, 79–103.
- Klein, A. M., Becker, E. S., & Rinck, M. (2011). Approach and Avoidance Tendencies in Spider Fearful Children: The Approach-Avoidance Task. *Journal of Child and Family Studies*, 20, 224–231.
- Koster, E. H. W., Crombez, G., Verschuere, B., van Damme, S., & Wiersema, J. R. (2006). Components of Attentional Bias to Threat in High Trait Anxiety: Facilitated Engagement, Impaired Disengagement, and Attentional Avoidance. *Behaviour Research and Therapy*, 44, 1757–1771.
- Kozlik, J., Neumann, R., & Lozo, L. (2015). Contrasting Motivational Orientation and Evaluative Coding Accounts: on the Need to Differentiate the Effectors of Approach/Avoidance Responses. *Frontiers in Psychology*, 6, 563.
- Kyrios, M., Frost, R. O., & Steketee, G. (2004). Cognitions in Compulsive Buying and Acquisition. *Cognitive Therapy and Research*, 28(2), 241–258.
- Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical Characteristics of Words Used in Emotional Stroop Experiments. *Emotion*, 6, 62–72.
- Lee, C.-C., & Chiou, W.-B. (2013). Keep Logging in! Experimental Evidence Showing the Relation of Affiliation Needs to the Idea of Online Social Networking. *Cyberpsychology, Behavior and Social Networking*, 16, 419–422.
- Lee, J., Bhatt, S., & Suri, R. (2018). When Consumers Penalize not so Green Products. *Psychology & Marketing*, 35, 36–46.
- Lee, S., Ha, S., & Widdows, R. (2011). Consumer Responses to High-Technology Products: Product Attributes, Cognition, and Emotions. *Journal of Business Research*, 64(11), 1195–1200.
- Lender, A., Meule, A., Rinck, M., Brockmeyer, T., & Bleichert, J. (2018). Measurement of Food-Related Approach-Avoidance Biases: Larger Biases When Food Stimuli are Task Relevant. *Appetite*, 125, 42–47.
- Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65, 272–292.
- Long, D. L., & Prat, C. S. (2002). Working Memory and Stroop Interference: An Individual Differences Investigation. *Memory & Cognition*, 30(2), 294–301.
- Luchs, M. G., Naylor, R. W., Irwin, J. R., & Raghunathan, R. (2010). The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference. *Journal of Marketing*, 74(5), 18–31.
- Lumsden, J., Skinner, A., Woods, A. T., Lawrence, N. S., & Munafò, M. (2016). The Effects of Gamelike Features and Test Location on Cognitive Test Performance and Participant Enjoyment. *PeerJ*, 4, e2184, 1–19.
- Luna, D., Ringberg, T., & Peracchio, L. A. (2008). One Individual, two Identities: Frame Switching among Biculturals. *Journal of Consumer Research*, 35, 279–293.
- Lusher, J., Chandler, C., & Ball, D. (2004). Alcohol Dependence and the Alcohol Stroop Paradigm: Evidence and Issues. *Drug and Alcohol Dependence*, 75, 225–231.
- MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: an Integrative Review. *Psychological Bulletin*, 109(2), 163–203.
- Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive Validity of the Implicit Association Test in Studies of Brands, Consumer Attitudes, and Behavior. *Journal of Consumer Psychology*, 14, 405–415.
- McKenna, F. P., & Sharma, D. (1995). Intrusive Cognitions: An Investigation of the Emotional Stroop Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 1595–1607.
- Mierke, J., & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, 85, 1180–1192.
- Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye Movements to Smoking-Related Pictures in Smokers: Relationship between Attentional Biases and Implicit and Explicit Measures of Stimulus Valence. *Addiction*, 98(6), 825–836.
- Munafò, M., Mogg, K., Roberts, S., Bradley, B. P., & Murphy, M. (2003). Selective Processing of Smoking-Related Cues in Current Smokers, Ex-Smokers and Never-Smokers on the Modified Stroop Task. *Journal of Psychopharmacology*, 17, 310–316.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, 19(6), 625–666.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. *Personality & Social Psychology Bulletin*, 31, 166–180.
- Ostafin, B. D., & Palfai, T. P. (2006). Compelled to Consume: The Implicit Association Test and Automatic Alcohol Motivation. *Psychology of Addictive Behaviors*, 20(3), 322–327.
- Payne, B. K., & Gawronski, B. (2010). A History of Implicit Social Cognition: Where Is It Coming From? Where Is It Now? Where Is It Going? In B. K. Payne & B. Gawronski (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications*, New York: Guilford Press, 1–15.
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An Inkblot for Attitudes: Affect Misattribution as Implicit Measurement. *Journal of Personality and Social Psychology*, 89(3), 277–293.
- Phillips, L. H., Bull, R., Adams, E., & Fraser, L. (2002). Positive Mood and Executive Function: Evidence from Stroop and Fluency Tasks. *Emotion*, 2, 12–22.
- Piqueras-Fiszman, B., Kraus, A. A., & Spence, C. (2014). "Yummy" versus "Yucky"! Explicit and Implicit Approach-Avoidance Motivations towards Appealing and Disgusting Foods. *Appetite*, 78, 193–202.
- Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.

- Ratcliff, R. (1993). Methods for Dealing with Reaction Time Outliers. *Psychological Bulletin*, 114, 510–532.
- Remedios, J. D., Chasteen, A. L., Rule, N. O., & Plaks, J. E. (2011). Impressions at the Intersection of Ambiguous and Obvious Social Categories: Does Gay+ Black= Likable? *Journal of Experimental Social Psychology*, 47(6), 1312–1315.
- Rinck, M., & Becker, E. S. (2007). Approach and Avoidance in Fear of Spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, 38(2), 105–120.
- Robinson, T. E., & Berridge, K. C. (2001). Incentive-Sensitization and Addiction. *Addiction (Abingdon, England)*, 96, 103–114.
- Roelofs, K., Elzinga, B. M., & Rotteveel, M. (2005). The Effects of Stress-Induced Cortisol Responses on Approach-Avoidance Behavior. *Psychoneuroendocrinology*, 30, 665–677.
- Roelofs, K., Putman, P., Schouten, S., Lange, W. G., Volman, I., & Rinck, M. (2010). Gaze Direction Differentially Affects Avoidance Tendencies to Happy and Angry Faces in Socially Anxious Individuals. *Behaviour Research and Therapy*, 48, 290–294.
- Rossiter, J. R. (2002). The C-OAR-SE Procedure for Scale Development in Marketing. *International Journal of Research in Marketing*, 19(4), 305–335.
- Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the Influence of Recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *The Quarterly Journal of Experimental Psychology*, 62(1), 84–98.
- Rozin, P., Hormes, J. M., Faith, M. S., & Wansink, B. (2012). Is Meat Male? A Quantitative Multimethod Framework to Establish Metaphoric Relationships. *Journal of Consumer Research*, 39, 629–643.
- Sachs, M., Kaplan, J., Der Sarkissian, A., & Habibi, A. (2017). Increased Engagement of the Cognitive Control Network Associated with Music Training in Children During an fMRI Stroop Task. *PloS one*, 12(10): e0187254, 1–29.
- Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology*, 8, p. 557.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual Differences in Components of Reaction time Distributions and their Relations to Working Memory and Intelligence. *Journal of Experimental Psychology. General*, 136, 414–429.
- Schnabel, K., Banse, R., & Asendorpf, J. (2006). Employing Automatic Approach and Avoidance Tendencies for the Assessment of Implicit Personality Self-Concept: The Implicit Association Procedure (IAP). *Experimental Psychology*, 53, 69–76.
- Schumacher, S. E., Kemps, E., & Tiggemann, M. (2016). Bias Modification Training Can Alter Approach Bias and Chocolate Consumption. *Appetite*, 96, 219–224.
- Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to Eat: How Immediate Affective and Motivational Responses to Food Cues are Influenced by Food Deprivation. *European Journal of Social Psychology*, 37(2), 359–379.
- Siegrist, M. (1997). Test-Retest Reliability of Different Versions of the Stroop Test. *The Journal of Psychology*, 131, 299–306.
- Simon, H. A. (1969). *The Sciences of the Artificial: 1st edition*. Cambridge, Mass: MIT Press.
- Solarz, A. K. (1960). Latency of Instrumental Responses as a Function of Compatibility with the Meaning of Eliciting Verbal Signs. *Journal of Experimental Psychology*, 59, 239–245.
- Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology*, 56(4), 283–294.
- Sriram, N., Nosek, B. A., & Greenwald, A. G. (2006). Scale Invariant Contrasts of Response Latency Distributions., Unpublished Manuscript, University of Virginia, Available at SSRN: <http://ssrn.com/abstract=2213910>.
- Stacy, A. W., & Wiers, R. W. (2010). Implicit Cognition and Addiction: a Tool for Explaining Paradoxical Behavior. *Annual Review of Clinical Psychology*, 6, 551–575.
- Steinhauser, M., & Hübner, R. (2009). Distinguishing Response Conflict and Task Conflict in the Stroop Task: Evidence from Ex-Gaussian Distribution Analysis. *Journal of Experimental Psychology. Human Perception and Performance*, 35, 1398–1412.
- Stormark, K. M., Field, N. P., Hugdahl, K., & Horowitz, M. (1997). Selective Processing of Visual Alcohol Cues in Abstinent Alcoholics: An Approach-Avoidance Conflict? *Addictive Behaviors*, 22, 509–519.
- Strauss, G. P., Allen, D. N., Jorgensen, M. L., & Cramer, S. L. (2005). Test-Retest Reliability of Standard and Emotional Stroop Tasks: an Investigation of Color-Word and Picture-Word Versions. *Assessment*, 12, 330–337.
- Stroop, J. R. (1935). Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Summers, T. A., & Hebert, P. R. (2001). Shedding some Light on Store Atmospherics: Influence of Illumination on Consumer Behavior. *Journal of Business Research*, 54(2), 145–150.
- Sun, R., Slusarz, P., & Terry, C. (2005). The Interaction of the Explicit and the Implicit in Skill Learning: a Dual-Process Approach. *Psychological Review*, 112, 159–192.
- Teichert, T., & Heidel, B. (2018). (Nicht-)Kompensatorik beim Kauf neuer FMCG-Produkte. *Transfer Werbeforschung & Praxis*, 64(1), 6–16.
- Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A Practical Guide to Implicit Association Tests and Related Tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York, New York: Guilford Press, 117–139.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions about Health, Wealth and Happiness*. London: Penguin Books Ltd.
- Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2008). Interactions between Implicit and Explicit Cognition and Working Memory Capacity in the Prediction of Alcohol Use in At-Risk Adolescents. *Drug and Alcohol Dependence*, 94, 116–124.
- van Holst, R. J., Lemmens, J. S., Valkenburg, P. M., Peter, J., Veltman, D. J., & Goudriaan, A. E. (2012). Attentional Bias and Disinhibition toward Gaming Cues are Related to Problem Gaming in Male Adolescents. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, 50, 541–546.
- Verbeken, S., Boendermaker, W. J., Loeys, T., Moens, E., Goossens, L., Latomme, J., & Braet, C. (2018). Feasibility and Effectiveness of Adding an Approach Avoidance Training with Game Elements to a Residential Childhood Obesity Treatment – A Pilot Study. *Behaviour Change*, 35(2), 91–107.
- Viswanathan, V., Sheppard, J. P., Kim, B. W., Plantz, C. L., Ying, H., Lee, M. J., Raman, K., Mulhem, F. J., Block, M. P., Calder, B., Lee, S., Mortensen, D. T., Blood, A. J., & Breiter, H. C. (2017). A Quantitative Relationship between Signal Detection in Attention and Approach/Avoidance Behavior. *Frontiers in Psychology*, 8, p. 122.
- Waters, A. J., & Feyerabend, C. (2000). Determinants and Effects of Attentional Bias in Smokers. *Psychology of Addictive Behaviors*, 14, 111–120.
- Waters, A. J., Sayette, M. A., Franken, I. H. A., & Schwartz, J. E. (2005). Generalizability of Carry-Over Effects in the Emotional Stroop Task. *Behaviour Research and Therapy*, 43, 715–732.
- Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive Bias Modification and Cognitive Control Training in Addiction and Related psychopathology. *Clinical Psychological Science*, 1, 192–212.
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining Automatic Action-Tendencies to Approach Alcohol in Hazardous Drinkers. *Addiction*, 105, 279–287.
- Wiers, R. W., Houben, K., Fadardi, J. S., van Beek, P., Rhemtulla, M., & Cox, W. M. (2015). Alcohol Cognitive Bias Modification Training for Problem Drinkers over the Web. *Addictive Behaviors*, 40, 21–26.

- Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The Emotional Stroop Task and Psychopathology. *Psychological Bulletin*, 120, 3–24.
- Woud, M. L., Becker, E. S., & Rinck, M. (2008). Implicit Evaluation Bias Induced by Approach and Avoidance. *Cognition & Emotion*, 22, 1187–1197.
- Woud, M. L., Maas, J., Wiers, R. W., Becker, E. S., & Rinck, M. (2016). Assessment of Tobacco-Related Approach and Attentional Biases in Smokers, Cravers, Ex-Smokers, and Non-Smokers. *Frontiers in psychology*, 7, p. 172.
- Yan, D. (2016). Numbers are Gendered: The Role of Numerical Precision. *Journal of Consumer Research*, 43, 303–316.
- Yu, Z., Wang, F., Wang, D., & Bastin, M. (2012). Beyond Reaction Times: Incorporating Mouse-Tracking Measures into the Implicit Association Test to Examine its Underlying Process. *Social Cognition*, 30, 289–306.
- Zajonc, R. B. (1968). Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology*, 9, 1–27.
- Zanto, T. P., Hennigan, K., Östberg, M., Clapp, W. C., & Gazzaley, A. (2010). Predictive Knowledge of Stimulus Relevance Does not Influence Top-Down Suppression of Irrelevant Information in Older Adults. *Cortex*, 46, 564–574.
- Znanewitz, J., Braun, L., Hensel, D., Altobelli, C. F., & Hattke, F. (2018). A Critical Comparison of Selected Implicit Measurement Methods. *Journal of Neuroscience, Psychology, and Economics*, 11, 249–266.

Keywords

Implicit Cognition, Marketing Research, Implicit Association Test, Emotional Stroop Task, Approach-Avoidance-Task