Caruelle, D., Gustafsson, A., Shams, P., & Lervik-Olsen, L. (2019). The use of electrodermal activity (EDA) measurement to understand consumer emotions–A literature review and a call for action. *Journal of Business Research*, *104*, 146-160.

https://doi.org/10.1016/j.jbusres.2019.06.041

The Use of Electrodermal Activity (EDA) Measurement to Understand Consumer Emotions

– A literature review and a call for action

Delphine Caruelle
PhD Candidate, Department of Marketing, BI Norwegian Business School, 0484 Oslo, Norway,
Email: delphine.caruelle@bi.no

Anders Gustafsson
Professor of Marketing, Department of Marketing, BI Norwegian Business School, 0484 Oslo, Norway
Email: anders.gustafsson@bi.no

Poja Shams
Lecturer in Business Administration, Karlstad University, 651 88 Karlstad, Sweden
Email: poja.shams@kau.se

Line Lervik-Olsen
Professor of Marketing, Department of Marketing, BI Norwegian Business School, 0484 Oslo, Norway
Email: line.lervik-olsen@bi.no

The Use of Electrodermal Activity (EDA) Measurement to Understand Consumer Emotions – A literature review and a call for action

Abstract

Electrodermal activity (EDA) is a psychophysiological indicator of emotional arousal. EDA measurement was first employed in consumer research in 1979 but has been scarcely used since. In the past decade, the ease of access to EDA recording equipment made EDA measurement more frequent in studies of consumer emotions. Additionally, recent calls to include physiological data in consumer studies have been voiced, which in turn is increasing the interest in EDA. Such a growing interest calls for assessing why and how EDA measurement has been used and should be used in consumer research. To this end, we undertook a critical review of studies of consumer emotions that employed EDA measurement. We found that most of these studies did not sufficiently report how they recorded and analyzed EDA data, which in turn impeded the replication of the findings. We therefore make recommendations derived from the psychophysiology literature to help consumer researchers get meaningful insights from EDA measurements. Finally, we call on researchers to be more transparent when reporting how they recorded and analyzed EDA data.

Keywords: electrodermal activity, customer experience, emotion, physiological measurement, arousal

The Use of Electrodermal Activity (EDA) Measurement to Understand Consumer Emotions – A literature review and a call for action

1. Introduction

We live in an experience-based economy and creating a strong customer experience is the main goal for many organizations (Lemon & Verhoef, 2016). It has been argued that customers value experiences over products and services (Abbott, 1955; Lemon & Verhoef, 2016). Consequently, instead of organizations competing on price, competing on customer experiences is an alternative route to take. Experiences prevail in customers' evaluations of brands because a brand can be defined as "a cluster of functional and emotional values that promises a unique and welcomed experience" (De Chernatony, 2010, p. 12). These emotional values are heavily influenced by smells, sounds, sights, tastes, and textures in any context. The question then is, how do organizations capture customer experiences and, in particular, their emotional component? The answer is usually by self-reports, such as questionnaires and other forms of verbal descriptions.

Self-reports, however, capture only one part of an emotional experience, which can be classified as the subjective part (Mauss & Robinson, 2009). An emotional experience does not simply consist of the description one can make of this experience. Apart from the subjective part, the experience also consists of a behavioral response (e.g., smiling) and of physiological changes (Bradley & Lang, 1999; Gross, 1998; Kring & Gordon, 1998; Lang, 1993; Mauss & Robinson, 2009). Physiological changes are indeed what differentiates affect from cognition (Lazarus, 1991). Emotions are first physiologically experienced (e.g., heart rate acceleration, sweat secretion); only as the mind perceives these physiological changes does the individual become aware of his/her emotion (Damasio, 2000). A consumer may never mentally be aware of the

physiologically change. Technological and neuroscientific advances, however, allow the unveiling of consumer emotions by studying physiological changes (i.e., reactions in the body). Of particular relevance to emotion research is electrodermal activity (EDA). EDA is a psychophysiological indicator of emotional arousal (Bradley & Lang, 1999; Critchley, 2002). As such, it enables researchers to measure emotions as they are produced through reactions in the body generated by brain signals (Sequeira, Hot, Silvert, & Delplanque, 2009). With EDA measurement, researchers obtain real-time data on consumers' emotional state, captured without any verbalization.

Consumer research can greatly benefit from measuring EDA to index an emotional response – provided that EDA measurement is appropriately employed. With the ease of access to EDA recording equipment, the interest in and use of EDA measurement have grown, which creates a need to better understand how this measurement method might help advance consumer research. However, despite EDA measurement having been used in consumer studies for 40 years, to the best of our knowledge, no research has been undertaken to assess and reflect upon the use of EDA measurement to understand consumer emotions in research. The objective of the present article is to fill this gap by offering a critical review of studies of consumer emotions that employ EDA measurement. To this end, we analyze why and how EDA measurement has been applied in the literature. Our review of 27 studies shows that an EDA measurement is used either alone or in addition to other measurement methods to obtain a more complete understanding of the emotional response. While the motivations for using EDA measurement are made clear, the method descriptions in these studies are to some degree scant, making replication slightly difficult. The majority of the studies in the review did not report, or inefficiently communicated, how EDA data were analyzed. This is the major flaw of previous studies that our review points

out. Consequently, we provide recommendations for analyzing EDA data and reporting their analysis.

This paper responds to calls from De Keyser, Lemon, Klaus, and Keiningham (2015), Lemon and Verhoef (2016), and Morales, Amir, and Lee (2017) to expand the scope of consumer research measurements to new types of data, including physiological data. To succeed in employing physiological measures such as EDA, consumer researchers need a common understanding of when and how to use such measures. We contribute to it by exploring the role of EDA measurement in consumer emotion research and by assessing the practical application in light of the psychophysiology literature.

The present paper is organized as follows. We first introduce EDA and detail the reasons for using EDA measurement. Next, we review studies of consumer emotions that employ EDA measurement by first identifying the relevant studies and briefly presenting them. Second, we examine the motives for using EDA measurement that these studies put forward and discuss the limitations and challenges they point out. Third, we examine the recording and analysis of EDA data in these studies and punctuate this examination with recommendations derived from the psychophysiology literature.

2. Basics of EDA

EDA refers to "electrical phenomena in skin" (Boucsein, 2012, p. 2) that can have a psychological significance. When an emotionally arousing stimulus is experienced, eccrine sweat glands produce sweat, which is an efficient conductor of current (Stern, Ray, & Quigley, 2000). As a result, the electrical properties of the skin change. The more emotionally arousing the stimulus is, the more sweat is secreted and the more the electrical properties of the skin change (for further details, see the Appendix on the psychological significance of EDA). Measuring EDA consists of measuring the electrical conductance, resistance, impedance, or admittance

(depending on the recording method) of the skin (Boucsein, 2012). Most often, skin conductance is measured and expressed in microsiemens (μS) (Stern et al., 2000).

EDA can be decomposed as tonic and phasic activity (Boucsein et al., 2012; Larkin, 2006; Stern et al., 2000). Figure 1 shows the decomposition of an EDA signal into its tonic and phasic components. Tonic activity corresponds to the "background" level of EDA and varies slowly. Tonic activity is referred to as the skin conductance level (SCL). By contrast, phasic activity corresponds to the response to a specific and discrete stimulus – such as an emotionally arousing one. When a stimulus, be it an object, a person, a situation, or a thought, is perceived as personally significant and gives rise to an emotional response (emotional arousal), the brain sends a signal through the sympathetic branch of the autonomic nervous system to the eccrine sweat glands in order to activate them (Dawson, Schell, & Filion, 2007; Larkin, 2006). These glands secrete sweat, which translates into a sudden increase in skin conductance; as a result, the skin conductance value increases. The entire process occurs within a few seconds: The increase in skin conductance starts 1 to 4 seconds after stimulus exposure, and it persists for 1 to 3 seconds (Boucsein et al., 2012; Dawson et al., 2007; Dawson, Schell, & Courtney, 2011). This sudden increase in skin conductance is referred to as skin conductance response (SCR) or, alternatively, as electrodermal activity response (EDR), or as peak. In the past, an SCR was commonly termed galvanic skin response (GSR).

---- Figure 1 about here ----

2.1 Reasons for employing EDA measurement

EDA measurement helps overcome three limitations inherent to self-reports of emotions: (1) the difficulty of obtaining a continuous measurement, (2) respondents' inability and/or

unwillingness to accurately report their emotions, and (3) the impossibility of capturing unconscious emotions.

First, self-reports can be taken only at certain points in time; they are usually taken before and/or after an event of interest to the researchers. Thus it is difficult, if not impossible, to obtain a continuous measurement of emotion with self-reports. By contrast, EDA allows capturing the emotional reaction continuously (Dawson et al., 2007). Some studies (e.g., Baumgartner, Sujan, & Padgett, 1997; Dalakas, 2006; Fredrickson & Levenson, 1998) seeking to obtain a continuous measurement of participants' emotions have used a cursor that participants had to move every time they felt a change in their emotional state. Nevertheless, such a tool does not guarantee a truly continuous measurement, because participants need to attend to their emotions and to think about reporting them before adjusting the cursor. In addition, the intrusiveness of the tool might affect the experience itself, and such effect might in turn bias the emotional measurement. In sum, to be truly continuous, a measurement of emotions cannot require individuals to report their emotions. EDA recording equipment is able to sample the electrical properties of the skin over time, making the measurement as fine grained and continuous as is technically feasible. Obtaining a continuous measurement is especially useful to identify which particular elements of a marketing stimulus generate arousal (LaBarbera & Tucciarone, 1995). EDA measurement is thus particularly relevant when the goal of the study is not compatible with a single account of arousal.

Second, social desirability bias can impede respondents from reporting their true emotions (Plassmann, Venkatraman, Huettel, & Yoon, 2015; Poels & Dewitte, 2006); and the recall of an emotion can differ from its actual experience (Alba & Williams, 2013; Cowley, 2008; Galak & Meyvis, 2011). By contrast, using EDA measurement allows bypassing the biases of the mind.

EDA is a physiological process controlled by the autonomic nervous system and, as such, it does not depend on voluntary control (Sequeira et al., 2009). Consumers might not always be able or willing to accurately tell about their emotions, whereas EDA captures the emotional response at the root (i.e., before any conscious cognitive processing) (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). In advertising research, copy testing is often subject to such biases, which is why EDA measurement is favored by some practitioners in this field (LaBarbera & Tucciarone, 1995).

Third, EDA measurement taps into unconscious emotions, which are emotions experienced physiologically but not reaching the stage of awareness (Critchley, 2002). An emotional experience is composed of a first level – the physiological changes – and of a second level – the awareness of these changes and, consequently, the awareness of the emotion (Wiens, 2005). In some instances, the second level does not occur, implying that the individual is not aware of the physiological changes happening in the body. This is termed an "unconscious emotion" (Winkielman & Berridge, 2004, p. 120). EDA measurement would capture the physiological changes, but the individual would not report feeling aroused. Importantly, these unconscious emotions affect behavior (Winkielman & Berridge, 2004), making them worth measuring.

While EDA measurement offers important advantages, it also raises limitations and challenges. These limitations and challenges must be considered before deciding to employ such measurement. First, EDA measurement does not provide any insight about the emotional valence; it simply indicates arousal level. Second, there is a possible risk of confounding with other mental processes when employing EDA measurement to index emotional arousal. Third, it is difficult to a posteriori identify the arousing stimulus that generated an emotional response, due to the 1-to-4-second latency between stimulus exposure and SCR onset. Fourth, the EDA

measurement method has been rapidly evolving, which requires researchers to constantly update their knowledge of the method. We discuss these limitations and challenges and how previous studies have addressed them later in the article.

3. Review of studies of consumer emotions that employ EDA measurement

To undertake this critical review, we needed to identify the relevant literature. As our focus is on studies of consumer emotions that employ EDA measurement, the following three criteria had to be met in order to include an article in the review. First, the article had to report one or more empirical studies that employ EDA measurement. Second, the article had to be in the realm of marketing or consumer research. Third, EDA measurement had to be employed to index an emotional response. EDA measurement has been used for quantifying physiological arousal in general, with no intent to specifically tap into emotional arousal (e.g., Alexander, Tripp, & Zak, 2015; Bradley, Angelini, & Lee, 2007; Kilbourne, Painton, & Ridley, 1985), or as an indicator of attention (e.g., Bolls, Muehling, & Yoon, 2003) or of implicit memory (e.g., Topolinski, Lindner, & Freudenberg, 2014). We disregarded such studies because we focus the discussion on the use of EDA as an indicator of emotional response.

We started the literature search in online databases. The keywords we used were "electrodermal activity" or, alternatively, other names given to this phenomenon ("skin conductance," "skin resistance," and "galvanic skin response") combined with the keyword "marketing" or alternatively "consumer." This search strategy proved to be of little efficiency: Most articles found in the search evoked EDA without reporting an empirical study employing its measurement (the first criterion was not met). We therefore adopted another search strategy – chain search (Rienecker, Jørgensen, & Skov, 2013). This strategy consists in "finding suitable literature by letting one text lead you to the next, which leads to the next and so on" (Rienecker et al., 2013, p. 148). Thus, we used as a starting point the articles we had already identified as

meeting all three criteria and searched for any other study that was cited in these articles and that seemed relevant to our review. We read these studies and included them in the review whenever they met all three criteria. We repeated the search process with every article newly identified as relevant literature.

This search strategy bears the risk of ignoring very recent studies since articles can cite past research only. To limit this risk, we conducted a similar search "forwards"; for each paper identified as relevant literature, we consulted on Google Scholar the list of articles citing it. As a result of this search, 27 articles were identified as relevant literature for our review of studies of consumer emotions that employ EDA measurement. We present them in the next section.

3.1 Overview of the studies in the literature

The use of EDA measurement in consumer research can be traced back to the seminal work by Kroeber-Riel in 1979, which reported EDA measurements from two unpublished dissertation works. Since then, multiple studies in advertising research have measured EDA to indicate emotional arousal level (e.g., Aaker, Stayman, & Hagerty, 1986; Bolls, Lang, & Potter, 2001; Droulers, Lajante, & Lacoste-Badie, 2013; Gakhal & Senior, 2008; Lajante, Droulers, Dondaine, & Amarantini, 2012; Micu & Plummer, 2010; Ohme, Reykowska, Wiener, & Choromanska, 2009; Vanden Abeele & MacLachlan, 1994; Venkatraman et al., 2015).

Furthermore, EDA measurement has been employed to indicate the level of emotional arousal elicited by a service recovery situation (Boshoff, 2012, 2017), by packaging (Hurley, Hutcherson, Tonkin, Dailey, & Rice, 2015), by touristic activities (Kim & Fesenmaier, 2015; Shoval, Schvimer, & Tamir, 2018), by a brand (Langner, Schmidt, & Fischer, 2015; Maxian, Bradley, Wise, & Toulouse, 2013; Reimann, Castaño, Zaichkowsky, & Bechara, 2012; Walla, Brenner, & Koller, 2011; Walla, Koller, Brenner, & Bosshard, 2017), by retail auctions (Adam,

Krämer, & Müller, 2015), by consumer reviews (Fox, Deitz, Royne, & Fox, 2018), and by shopping environments (Groeppel-Klein, 2005; Groeppel-Klein & Baun, 2001). EDA measurement can also be found in studies of the impact of arousal on attitude toward a product (Bettiga, Lamberti, & Noci, 2017) and on purchase decisions (Guerreiro, Rita, & Trigueiros, 2015; Somervuori & Ravaja, 2013). Table 1 provides a summary of these studies of consumer emotions that employ EDA measurement.

---- Table 1 about here ----

This review brought to our attention theoretical and practical considerations on implementing a study with EDA measurement. We address these considerations in the rest of the paper.

4. Theoretical considerations when employing EDA measurement

In this section, we discuss two major factors that marketing researchers must consider before implementing a consumer study that uses EDA measurement: What are the motives for employing EDA measurement? And what are the limitations and challenges to be aware of when employing EDA measurement?

We have identified five motives for using EDA measurement in studies of consumer emotions: *exploring and assessing this measurement method, avoiding cognitive biases, tapping into unconscious processes, obtaining a real-time measurement,* and *capturing multiple dimensions of the emotional response*. These motives are not necessarily mutually exclusive and there can be multiple motives for using EDA measurement in a study of consumer emotions. In the following discussion, we focus on what we have identified as the primary motive in each study.

First, for 5 of the 27 studies (19%) listed in Table 1, the motive for using EDA measurement was to explore and assess this measurement method. This is evident in the work by

Kroeber-Riel (1979), who pioneered the use of EDA measurement in consumer research. Similarly, the studies by Lajante et al. (2012) and by Droulers et al. (2013) were driven by the willingness to explore the method. Micu and Plummer (2010) employed physiological measurements (including EDA) to compare the insights they provide with insights from self-reports. Similarly, Venkatraman et al. (2015) employed EDA measurement together with other measurement methods (e.g., fMRI, eye tracking) to determine which measurement is best at predicting the market-level advertising elasticities.

Second, in another 5 of the 27 studies (19%) listed in Table 1, EDA measurement was employed to avoid the cognitive biases inherent to self-reports (e.g., Boshoff, 2012, 2017; Guerreiro et al., 2015; Somervuori & Ravaja, 2013; Walla et al., 2011). Some study contexts are particularly likely to make respondents unwilling to reveal their true emotions. For instance, Boshoff (2012) investigated the consumer response to service employees of the same or a different ethnicity in a service recovery situation. As emphasized by Boshoff, social desirability bias is likely to emerge when studying ethnicity, making self-reports unreliable. For this reason, Boshoff used EDA measurement to quantify the degree of emotional arousal generated by an interaction with an employee of the same or a different ethnicity.

Third, EDA measurement was used to tap into unconscious processes in 2 of the 27 studies (7%) listed in Table 1. Ohme et al. (2009) aimed to compare the emotional response to two versions of a scene in a TV commercial, two versions that participants could not consciously discern, making EDA measurement particularly useful to tap into the unconscious emotional response. Furthermore, Gakhal and Senior (2008) sought to investigate the involvement of the right hemisphere of the brain in processing advertisements showing famous and/or beautiful models, and opted for EDA measurement given the assumption that one hemisphere of the brain activates EDA on the opposite side of the body.

Fourth, in 4 of the 27 studies (15%) reported in Table 1, recording EDA was performed to obtain a real-time measurement of the emotional response. This simultaneity was sought for by researchers who investigated the emotional experience during extended episodes (e.g., Groeppel-Klein, 2005; Groeppel-Klein & Baun, 2001; Kim & Fesenmaier, 2015; Shoval et al., 2018). Participants in a study involving the experience of an extended episode might have difficulties remembering how they felt throughout the extended episode (Groeppel-Klein, 2005; Groeppel-Klein & Baun, 2001). Thus, a retrospective measurement (such as a self-report) is not well-suited. By contrast, EDA measurement is particularly appropriate because it measures the emotional response in real time, as it occurs. Furthermore, another reason why EDA measurement is useful for extended episodes is that it allows tracing the emotional response over time, as the extended episode unfolds (e.g., Kim & Fesenmaier, 2015; Shoval et al., 2018).

Fifth, in 11 of the 27 studies (41%) listed in Table 1, EDA measurement was combined with other measurements (e.g., self-reports) to capture multiple dimensions of the emotional response. Examples of studies with such a motive are Aaker et al. (1986), Adam et al. (2015), Bettiga et al. (2017), Bolls et al. (2001), Fox et al. (2018), Hurley et al. (2015), Langner et al. (2015), Maxian et al. (2013), Reimann et al. (2012), Vanden Abeele and MacLachlan (1994), and Walla et al. (2017). Such combining is in line with the recommendation by Hirschman and Holbrook (1982), De Keyser et al. (2015), Lemon and Verhoef (2016), and Verhoef et al. (2009) to embrace a multi-method approach, including physiological measurements, to fully comprehend the consumers' emotional response and the richness of the customer experience. Some studies that combined EDA measurement with self-reports of arousal found corroborating results (e.g., Adam et al., 2015; Bolls et al., 2001; Langner et al., 2015; Reimann et al., 2012), while others did not (e.g., Bettiga et al., 2017; Maxian et al., 2013; Walla et al., 2017).

To conclude this review of the motives for employing EDA measurement, we would like to point out that EDA should not be seen as a perfect substitute for self-reports. Rather, the context and the purpose of the study should guide researchers in determining what type of measurement suits best. In light of the properties of EDA measurement discussed in Section 2, we can make the following recommendations. Self-reports can be employed when there is little doubt that participants in the study are able and willing to express their emotions. However, if the study relates to a sensitive topic, EDA measurement is better suited because it cannot be distorted by social desirability bias (Poels & Dewitte, 2006). Furthermore, if the study aims to investigate how the emotional response varies over time, EDA measurement is particularly appropriate given the continuous trace obtained with EDA recording equipment (Dawson et al., 2007). But, if only one discrete measurement is needed, a self-report can suffice. Importantly, EDA measurements and self-reports can be used conjointly in a study, as complements. If researchers have the opportunity to employ both methods, they can both record EDA and administer a self-report to get a more complete overview of the emotional response.

4.1 Limitations of and challenges raised by EDA measurement

In our review of the studies listed in Table 1, we have identified four limitations and challenges associated with using EDA measurement: *lack of insight about the emotional valence*, *possible confounding with other mental processes, difficulty to identify the arousing stimulus a posteriori*, and *rapid evolution of the method*. First, as pointed out by Groeppel-Klein and Baun (2001) and Groeppel-Klein (2005), EDA measurement does not offer insights about the valence of the emotional experience; it simply indicates arousal level. An emotional experience is characterized by its valence and by its arousal level, the combination of which form discrete emotional states such as sadness or delight (Mauss & Robinson, 2009; Russell, 1980). Thus, EDA measurement can be used alone in studies focusing on the arousal dimension of the

emotional experience, but needs to be supplemented with other emotion measurements if researchers seek to determine its valence as well. There are multiple ways to supplement EDA measurement. A self-report can be administered to identify the valence of the emotional experience. For example, Guerreiro et al. (2015) used, in addition to EDA measurement, the self-assessment manikin to determine the degree of pleasure or displeasure induced by different products. Similarly, Groeppel-Klein (2005) and Shoval et al. (2018) used verbal scales to determine the valence of the consumer experience. Another way to assess the valence or type of emotion is to conduct post-study interviews, as evidenced by Hurley et al. (2015) and Kim and Fesenmaier (2015). However, when employing self-reports or interviews in addition to EDA measurements, the researcher reintroduces some of the issues (e.g., verbalization and recall biases) that the use of EDA measurement has helped overcome. To avoid this drawback, EDA measurements can be complemented with other physiological measurements, such as startle reflex modulation (e.g., Walla et al., 2017), facial electromyography (e.g., Somervuori & Ravaja, 2013), and/or electroencephalography (e.g., Boshoff, 2012, 2017; Ohme et al., 2009). Startle reflex modulation measures the amplitude of an eyeblink as the response to a startle probe – a decreased amplitude indicating a positive emotion and an increased amplitude indicating a negative emotion. Facial electromyography records facial muscle activity: A frown indicates a negative emotion whereas a smile indicates a positive emotion. Electroencephalography can help determine the valence of the emotional experience, with left-frontal activity being associated with positive emotions and right-frontal activity being associated with negative emotions.

A second challenge, as noted by Kroeber-Riel (1979), is that EDA indicates not only emotional arousal, but also attention. In addition, it can indicate cognitive processing (Dawson et al., 2007; Stern et al., 2000). Therefore, studies of emotional arousal that use EDA measurement must be designed to control for these other processes. For instance, when seeking to compare the

emotional response to two different advertisements, researchers must ensure that these two advertisements require the same amount of cognitive processing. To this end, the study by Witt, reported by Kroeber-Riel (1979), used similar advertisements, keeping the slogan, copy, and brand names identical, just modifying the illustration (a person wearing more or less clothing), thus ensuring that the difference in EDR is driven solely by the change in illustration and not by a copy that would be more challenging to process.

A third challenge with EDA data, as stressed by Hurley et al. (2015), is the difficulty to a posteriori identify the arousing stimulus that generated a response. This difficulty is due to the 1-to-4-second latency between the exposure to an arousing stimulus and the change in EDA level (an SCR) that it generates (Boucsein et al., 2012; Dawson et al., 2007). This challenge is particularly salient when multiple stimuli were presented within a short time interval. For instance, in the study by Hurley et al. (2015), which sought to investigate the emotional response to packaging, participants were free to browse in a fictitious shopping environment. As participants' gaze quickly shifted from one packaging to another, the authors had difficulty associating an SCR to a gaze at a specific packaging. To avoid such difficulty, researchers can design the study so as to sufficiently space out (over time) the stimuli they expose participants to. For example, in the study by Gakhal and Senior (2008), participants saw a blank screen for 10 seconds between each presentation of the stimulus.

A fourth challenge we have identified is the rapid evolution of EDA recording methods. Early consumer research studies employing EDA measurement (e.g., Aaker et al., 1986; Kroeber-Riel, 1979) reported using a polygraph that measured skin resistance. However, more recent devices measure skin conductance (e.g., Affectiva Q-Sensor, Biopac BioNomadix, Coulbourn Instruments skin conductance coupler, Empatica E4). The distinction is critical because skin

conductance data and skin resistance data require different analyses (Boucsein, 2012; Stern et al., 2000). In addition, the polygraph's output was on paper, whereas current devices offer digital storage and display. Researchers must therefore pay particular attention when reading past EDA studies to discern which information is still relevant given the rapid technological changes. Nevertheless, it should be noted that the rapid evolution of EDA recording methods also constitutes an opportunity. For instance, the size of the recording equipment had previously complicated field studies. Groeppel-Klein (2005, Studies 3 and 4) needed to ask participants to carry a shoulder bag in which the equipment was placed. Since then, wireless wristbands (e.g., Affectiva Q-Sensor, Biopac BioNomadix, Empatica E4) have been made available, which facilitates ambulatory recordings as illustrated in the studies by Kim and Fesenmaier (2015), Hurley et al. (2015), and Shoval et al. (2018).

5. Practical considerations when employing EDA measurement

In this section, we address the practices related to recording and analyzing EDA data in consumer studies. A summary of the principal practices reported in the aforementioned studies is provided in Table 2.

--- Table 2 about here ----

5.1. Recording EDA data

Recording EDA requires EDA recording equipment. Not all equipment uses the same recording technique (Boucsein, 2012). The recording can be exosomatic or endosomatic, depending on whether an external voltage is applied to the skin (exosomatic recording) or not (endosomatic recording). Exosomatic recording is done by applying a direct current (DC) or an alternating current (AC). Furthermore, some EDA recording equipment allows recording phasic EDA separately from tonic EDA, with the help of an AC-coupled amplifier (Boucsein, 2012;

Boucsein et al., 2012; Stern et al., 2000). Only 6 of the 27 studies (22%) listed in Table 2 explicitly reported which recording technique they used. All studies that did so reported using the exosomatic DC technique (Adam et al., 2015; Groeppel-Klein, 2005; Groeppel-Klein & Baun, 2001; Guerreiro et al., 2015; Lajante et al., 2012; Somervuori & Ravaja, 2013). For the other studies, even though they did not explicitly indicate which recording technique they used, the fact that they reported conductance or resistance data suggests they have performed an exosomatic recording with DC or with AC converted to DC. Only one study (Boshoff, 2017) does not provide any information that would allow deducing the type of EDA recording.

The exosomatic recording technique consists of passing current across the skin, between two electrodes that are part of the EDA recording equipment. There are multiple locations on the body where it is possible to place these electrodes (Boucsein, 2012; van Dooren, de Vries, & Janssen, 2012). Frequently used locations include the fingers (e.g., Aaker et al., 1986; Boshoff, 2012; Fox et al., 2018; Gakhal & Senior, 2008; Guerreiro et al., 2015; Lajante et al., 2012; Ohme et al., 2009; Somervuori & Ravaja, 2013; Venkatraman et al., 2015; Walla et al., 2011; Walla et al., 2017), the palms (e.g., Adam et al., 2015; Bolls et al., 2001; Groeppel-Klein & Baun, 2001; Langner et al., 2015; Vanden Abeele & MacLachlan, 1994), and the wrists (e.g., Bettiga et al., 2017; Kim & Fesenmaier, 2015; Hurley et al., 2015). When possible, fingers and palms should be preferred to other locations, because they have a high density of eccrine sweat glands (Boucsein et al., 2012; van Dooren et al., 2012). Soles of the feet and shoulders have been shown to be good alternatives to the palms and fingers, whereas upper arms, backs, and armpits should be avoided (van Dooren et al., 2012). It should be noted that some EDA recording devices are conceived for recording EDA at one location only. Thus, the choice of where to place the electrodes for recording EDA is sometimes tied to the choice of the device.

Typically, when the electrodes are placed on the fingers, palms, or wrists, the non-dominant hand is equipped with the electrodes of the EDA recording equipment (e.g., Aaker et al., 1986; Adam et al., 2015; Droulers et al., 2013; Guerreiro et al., 2015; Groeppel-Klein & Baun, 2001; Hurley et al., 2015; Lajante et al., 2012; Langner et al., 2015; Maxian et al., 2013; Somervuori & Ravaja, 2013; Vanden Abeele & MacLachlan, 1994; Walla et al., 2011; Walla et al., 2017). Equipping the non-dominant hand allows limiting movement artifacts during EDA recording. As an additional preventive measure, researchers can explicitly instruct participants to avoid movements of the equipped hand – as Maxian et al. (2013) and Bettiga et al. (2017) did. Gakhal and Senior (2008) chose to equip both hands because they aimed to distinguish EDA recorded on the left hand (that would correspond to signals generated by the right hemisphere of the brain) from EDA recorded on the right hand (that would correspond to signals generated by the left hemisphere). However, such a practice is questionable because the lateralization of EDA – i.e., that one hemisphere of the brain activates EDA on the opposite side of the body – has not been firmly established (Boucsein, 2012; Dawson et al., 2007; Picard, Fedor, & Ayzenberg, 2016).

Bettiga et al. (2017) reported recording EDA during a resting period, before the actual study started, to serve as a baseline. Such a practice is not necessary according to Boucsein (2012), since phasic EDA values do not depend on tonic EDA values.

The manipulation of the EDA recording equipment can be challenging and can lead to excluding some EDA recordings from the analysis. For instance, Maxian et al. (2013) excluded the EDA recordings of six participants from the analysis because of "equipment error" (p. 472). Adam et al. (2015) reported that the EDA measurements "failed completely" (p. 474) for 16 participants and that 19 other EDA recordings had to be excluded from the analysis because of noise in the data or aberrant EDA values. Fox et al. (2018) evoked the exclusion of "outliers" (p.

49) from the analysis, though they did not indicate why those recordings were identified as outliers. Thus, researchers might want to increase the initial sample size in case technical problems with the equipment occur during EDA recording.

Walla et al. (2011) and Somervuori and Ravaja (2013) mentioned that participants in their studies were healthy individuals. This recruitment criterion relates to the fact that certain pathologies (e.g., schizophrenia, anxiety) can affect EDA (Boucsein, 2012). Therefore, it is recommended to record EDA of healthy individuals. Besides pathologies, demographics (age, gender, and ethnicity) can affect EDA (Boucsein et al., 2012). For instance, older adults tend to have lower skin conductance values than younger adults do (Boucsein, 2012). However, unless the purpose of the study is to investigate the emotional experience of a specific demographic group, demographic characteristics should not serve as recruitment criteria, because such criteria might considerably limit the external validity of the findings. Solutions to account for such inter-individual differences in skin conductance values are introduced in Section 5.4.

5.2. Preprocessing EDA data

To facilitate analyzing the EDA data, some transformations can be performed on the raw data. One such transformation is downsampling, reported by 3 of the 27 studies listed in Table 2 (11%). The EDA recording equipment samples EDA at a certain frequency – referred to as "sampling rate." Researchers can decide to downsample (i.e., to assign a lower sampling rate) if they wish to speed up the processing time – the lower the sampling rate, the less time it takes for the software program to process the data. Downsampling can be performed either during the recording (provided that the EDA recording equipment allows customizing of the sampling rate) or as an intermediate step between recording and analyzing the EDA data (in a software program).

Another transformation, reported by 30% of the studies listed in Table 2, consists of performing data smoothing (e.g., Droulers et al., 2013; Lajante et al., 2012; Hurley et al., 2015), or applying a low-pass filter on the EDA signal (e.g., Boshoff, 2012; Fox et al., 2018; Guerreiro et al., 2015; Langner et al., 2015; Ohme et al., 2009). This transformation is required to treat the high-frequency noise that the EDA recording equipment produces. In addition, smoothing or applying a low-pass filter to the raw EDA data helps to remove small artifacts caused by movements (Xia, Jaques, Taylor, Fedor, & Picard, 2015). This transformation can be performed either while recording (if the EDA recording equipment offers such a feature) or in a software program, before analyzing the data.

An additional transformation is artifact removal. Smoothing or applying a low-pass filter to the EDA signal is not always sufficient to remove all artifacts (Xia et al., 2015). Despite precautions taken when recording EDA (e.g., equipping the non-dominant hand), artifacts might still occur, in particular when EDA is recorded in ambulatory settings. For instance, a participant pulls a produce bag in a store, which results in an artifact in the EDA data (Groeppel-Klein & Baun, 2001). In any case, it is particularly important to deal with artifacts because the phasic EDA detection might mistake them for SCRs.

Lajante et al. (2012) and Bettiga et al. (2017) conducted a visual inspection of the EDA signal in order to detect artifacts. Such a visual inspection requires knowing what an SCR "looks like." An SCR is characterized by a rising time lasting from 1 to 3 seconds (Dawson et al., 2007; Dawson et al., 2011) and the slope following the apex is less steep because it takes 2 to 10 seconds for a 50% recovery of the SCR amplitude. Because the declining slope is less steep than the rising slope, an SCR has an asymmetric shape (see Figure 2A). By contrast, a movement artifact occurs extremely fast and the "recovery" is almost immediate – the rising and the declining slopes both look very steep (see Figure 2B). The time frame and the shape that

characterize an SCR are thus different from those of a movement artifact. Similarly, a sudden drop in the EDA signal is likely to be due to a physical disconnection between the electrodes of the EDA recording equipment and the skin (see Figure 2C for an illustration). For instance, having the hand resting on a table has ensured a good connection between the skin and the electrodes that have been loosely attached to that hand. However, as soon as the participant moves the hand away from the table, this connection is lost, which makes the EDA value suddenly drop until the participant places the hand back on the table.

---- Figure 2 about here ----

Performing a visual inspection can be time consuming, especially if the EDA recordings are long and/or numerous. Indeed, in the early 2000s, Groeppel-Klein and Baun (2001) and Groeppel-Klein (2005) deplored the lack of an automated artifact detection program. It is now possible to automatically detect artifacts using a software program called EDA Explorer. This program employs an algorithm developed using machine learning (Taylor et al., 2015). It indicates which time windows (of 5 seconds each) contain artifacts.

Artifacts detected by visual or automated inspection must be removed, as documented in the studies by Groeppel-Klein and Baun (2001), Groeppel-Klein (2005), Lajante et al. (2012), Droulers et al. (2013), Kim and Fesenmaier (2015), and Bettiga et al. (2017). Removal can be done "manually" by deleting artifacts one by one: The EDA signal is modified so as to delete the data points corresponding to the artifact and is reconstructed by interpolation. If artifacts cannot be removed (e.g., too many and/or too long recordings for manual removal to be feasible), Boucsein (2012) recommends disregarding the time windows in which artifacts have been detected.

In sum, even though low-pass filtering and artifact removal are generally recommended (Boucsein, 2012; Boucsein et al., 2012), no standard for preprocessing EDA data exists.

Therefore, we strongly encourage researchers to be transparent about how they proceeded and to report as completely as possible what transformations they performed on their raw EDA data.

5.3. Detecting and quantifying phasic EDA

Analyzing EDA data primarily consists of detecting and quantifying the phasic component of the EDA signal. Unless purposefully focusing on tonic EDA (e.g., Shoval et al., 2018), researchers should detect and quantify phasic EDA to index the level of arousal generated by a discrete stimulus. Surprisingly, of the 25 studies listed in Table 2 and for which this step is applicable, as many as 14 studies (56%) did not report how researchers proceeded to detect and quantify the phasic response. This is problematic because knowing how this detection was performed is crucial in assessing the validity of the findings and in ensuring their replicability. Not reporting the phasic EDA detection method could also mean that phasic EDA was recorded separately from tonic EDA (with the help of an AC-coupled amplifier as part of the EDA recording device). In such a case, for the sake of clarity, researchers should explicitly state that they have recorded phasic EDA separately from tonic EDA. For instance, Groeppel-Klein (2005) implied having done so when she mentioned having recorded EDR specifically. Please note that recording phasic EDA separately from tonic EDA enables researchers to skip the detection step but does not enable to ignore all quantification matters such as overlapping SCRs, as the discussion below will show.

Depending on the recording technique, skin conductance, resistance, impedance, or admittance is measured. Except for Aaker et al. (1986) and Kroeber-Riel (1979), who record skin resistance, and Boshoff (2017), who does not report the recording technique, all studies listed in Table 2 record skin conductance. Therefore, the following discussion relates specifically to the analysis of skin conductance data.

In four of the studies listed in Table 2, a baseline is used to quantify the response to a stimulus. The baseline approach, however, covers multiple ways to proceed. Somervuori and Ravaja (2013) calculated the baseline as the mean EDA value during the 5 seconds that precede the stimulus presentation and subtracted it from the mean EDA value during stimulus viewing. Maxian et al. (2013) computed the baseline similarly but, unlike Somervuori and Ravaja, they did not subtract it from the mean EDA value during stimulus viewing. Rather, they selected the highest EDA value during the 1 to 4 seconds following the stimulus onset and computed the change from the baseline value to the highest value. Gakhal and Senior (2008) reported a more complex way to proceed with a baseline, in 3 steps. First, they determined the EDA value during the middle third of the stimulus presentation. Then they added to this value the lowest EDA value recorded during a baseline (without indicating when this baseline was measured). Last, the highest EDA value (presumably from the last third of the stimulus presentation) was compared to the value computed in the two previous steps. Finally, Reimann et al. (2012) offered a different approach to the baseline, which they define as the minimum SCR that occurs between 0 and 4 seconds following the stimulus onset. They subtracted the value of this baseline from the "peak SCR" (p. 134) that occurs between 1 and 4 seconds after stimulus onset. How these SCRs were detected in the first place is, regrettably, not reported. Strikingly, none of the four studies mentioned above reported prior research that would justify the methods they used to detect and quantify the phasic response. We strongly encourage researchers to cite (when reporting the EDA data analysis) prior research that has validated the method used to detect and quantify the phasic response. This would help the reader make an informed assessment of the quality of the data analysis.

Besides using a baseline, a few other methods to detect the phasic response have been documented in consumer research. Ohme et al. (2009, p. 26) and Boshoff (2012, p. 404) reported

using “differential analysis, wavelet transformation, and other mathematical and statistical tools,” which leaves room for interpretation. Venkatraman et al. (2015) and Langner et al. (2015) detected phasic EDA with the help of Acqknowledge software’s analysis that performs high-pass filtering. Bettiga et al. (2017), Droulers et al. (2013), and Lajante et al. (2012) used the Ledalab software program, in which they performed continuous decomposition analysis – a method to detect phasic EDA developed by Benedek and Kaernbach (2010a).

Adam et al. (2015) and Hurley et al. (2015) reported the software programs they employed to detect phasic EDA but did not give any further information regarding the detection method they used. This lack of precision is regrettable because it does not allow readers to know which method was employed to detect phasic EDA. For instance, Adam et al. (2015) specified detecting phasic EDA with Ledalab; however, this software program offers three detection methods (continuous decomposition analysis, discrete decomposition analysis, and trough-to-peak analysis) (Benedek, n.d.).

To detect and quantify the phasic response, researchers must determine which time window is of interest. First, some studies seek to quantify the overall emotional response to an event that is deployed over time. This event can be deployed over half a minute (e.g., a TV commercial in Droulers et al., 2013; Lajante et al., 2012; Venkatraman et al., 2015), over several minutes (e.g., an interaction with a product in Bettiga et al., 2017; a store visit in Groeppel-Klein, 2005 and in Groeppel-Klein & Baun, 2001) or over a few hours (e.g., a touristic activity in Kim & Fesenmaier, 2015). Such events are likely to comprise multiple discrete stimuli that give rise to an emotional response. In this instance, the time window to analyze EDA data corresponds to the entire duration of the event.

Other studies are interested in quantifying the emotional response to a short-lasting, ephemeral stimulus. This stimulus can be the presentation of a logo (e.g., Maxian et al., 2013;

Reimann et al., 2012), the viewing of a short scene within a TV commercial (e.g., Ohme et al., 2009), or the communication of the outcome of an auction (e.g., Adam et al., 2015). The time window to detect and quantify the phasic response to such stimuli typically corresponds to the 1-to-4-second window after stimulus exposure, as this window equates the latency that separates an arousing stimulus from the onset of the SCR it generates (Boucsein et al., 2012; Dawson et al., 2007).

Depending on the software program used, and on the method selected, a certain number of criteria must be set in order to detect and quantify the phasic response. One usual criterion is the minimum amplitude for SCRs (i.e., a threshold that the rise in skin conductance value must reach or surpass to qualify as SCR). It is generally set at a value ranging from 0.01 to 0.05 µS (Boucsein et al., 2012). The recording conditions should guide researchers in determining the exact threshold (Boucsein et al., 2012). A threshold as low as 0.01 µS suffices for EDA data recorded in a controlled environment, with no artifacts or with artifacts removed as part of the preprocessing. This is the case for the studies by Adam et al. (2015) and by Lajante et al. (2012). These researchers set a criterion of 0.01 µS as the minimum amplitude for SCRs, which is in accordance with the conditions in which they recorded EDA: in a controlled lab environment (no movement by participants). However, a higher threshold is preferred if the EDA data were recorded in a less controlled environment, such as an ambulatory setting or a field study. For instance, Groeppel-Klein and Baun (2001) conducted a field study where participants were free to move in a store. Therefore, they considered only SCRs reaching a minimum amplitude of 0.05 µS.

Several challenges emerge with the detection and quantification of phasic EDA. A first challenge relates to overlapping SCRs (see Figure 3 for an illustration). SCRs are said to overlap when a new SCR occurs before the EDA signal has entirely recovered from the previous SCR

(Benedek & Kaernbach, 2010a; Boucsein et al., 2012). Overlapping SCRs are problematic because the onset of the newly occurring SCR is masked by the trail that remains from the previous SCR. This masking can generate two problems, one related to detecting SCRs (Droulers et al., 2013; Lajante et al., 2012) and one related to estimating the SCR amplitude (Groeppel-Klein & Baun, 2001; Groeppel-Klein, 2005).

---- Figure 3 about here ----

First, overlapping SCRs complicate the detection of SCRs. Certain methods traditionally used in psychophysiology to detect phasic EDA (e.g., trough-to-peak analysis) do a poor job of detecting SCRs when they are overlapping (Alexander et al., 2005; Benedek & Kaernach, 2010a, 2010b). Therefore, when overlapping SCRs are observed in the EDA signal, researchers must select a suitable method to detect the phasic response. For instance, the presence of overlapping SCRs in the EDA signal motivated Droulers et al. (2013) and Lajante et al. (2012) to use continuous decomposition analysis, which was developed to suit the analysis of EDA signals with overlapping SCRs (Benedek & Kaernbach, 2010a).

Second, overlapping SCRs complicate the quantification of the SCR amplitude. Once an SCR has been detected, its amplitude is quantified as the difference between the skin conductance value at the apex of the SCR and the skin conductance value at the onset (Stern et al., 2000). When SCRs are overlapping, this quantification procedure is complicated by the fact that the "true" skin conductance value at the onset is unknown, because the trail from the previous SCR is superimposed on the onset of the newly occurring SCR. Several solutions to quantify overlapping SCRs have been proposed (for a review, see Boucsein, 2012). Researchers should report which solution they selected, as Groeppel-Klein and Baun (2001) and Groeppel-Klein (2005) did.

A second challenge is habituation of the SCR (Boucsein et al., 2012; Dawson et al., 2007; Stern et al., 2000). Habituation refers to the fact that the response to a stimulus becomes smaller

and smaller as the individual is exposed to the same or to a similar stimulus again and again. SCRs are prone to such habituation: Their amplitudes diminish as the individual for whom EDA is measured is repeatedly exposed to the same stimulus. The quantification of EDA is thus altered by this phenomenon. Habituation is particularly likely to arise in studies with a repeated-measure design, because participants in such studies are exposed multiple times to the same stimulus. Therefore, researchers who employ a repeated-measure design need to be aware that the quantification of phasic EDA is affected by the repeated exposure of participants to the same stimulus. Among the studies listed in Table 2, only the study by Adam et al. (2015) tested – and found evidence – for habituation because each participant was exposed 15 times to the stimulus (being told the outcome of an auction).

5.4. Selecting EDA metrics

The outputs obtained from the EDA analysis are manifold; 26% of the studies listed in Table 2 reported multiple EDA metrics. First, the occurrence of an SCR is informative – because the non-occurrence of an SCR would indicate that the stimulus failed to generate a response (Boucsein et al., 2012). This is evidenced in the study by Hurley et al. (2015), in which no SCR was found for 33 of the 36 participants.

Second, and conditional on the occurrence of an SCR, the amplitude of the SCR (i.e., the number of μS by which skin conductance value rose) is a particularly relevant metric because it enables estimating the degree of arousal experienced by an individual (Dawson et al., 2007). Such a metric was embraced by Langner et al. (2015) and by Venkatraman et al. (2015). Multiple SCRs can occur in the time window chosen for analyzing EDA data. In this case, it is a common practice (19% of the studies listed in Table 2) to compute the sum of the amplitudes of the SCRs (e.g., Bettiga et al., 2017; Groeppel-Klein, 2005; Groeppel-Klein & Baun, 2001; Lajante et al., 2012) or the average SCR amplitude (e.g., Ohme et al., 2009).

Another relevant metric is the area under the curve, which accounts for the duration of the SCR, in addition to its amplitude (Boucsein et al., 2012). A variant of the area under the curve is ISCR (integrated SCR), proposed by Benedek and Kaernbach (2010a). ISCR is the area under the curve of phasic EDA within the response window that the researcher has specified (Benedek & Kaernbach, 2010a). ISCR is particularly suitable for extended time windows, in which multiple SCRs are likely to occur. Consumer studies in which ISCR was employed include those of Bettiga et al. (2017), Droulers et al. (2013), and Lajante et al. (2012).

Bolls et al. (2001), Groeppel-Klein (2005), and Groeppel-Klein and Baun (2001) calculated the number of SCRs for each participant's EDA recording and used this number as an indicator of the emotional response. This metric can be complemented with the average amplitude of the SCRs to obtain a more complete account of the emotional arousal level.

Guerreiro et al. (2015) employed SCR rise time and Langner et al. (2015) employed SCR rise time and recovery time. Nevertheless, the psychological meanings of SCR rise time and of SCR recovery time are not well established (Dawson et al., 2007). Thus, we would recommend against using them as indicators of emotional arousal levels.

Strikingly, 15 of the 27 studies (56%) listed in Table 2 failed to effectively communicate the EDA metrics they employed. For instance, Boshoff (2012, p. 406) indicated "galvanic skin response" as the EDA metric he opted for. However, galvanic skin response refers to the phenomenon (the change in skin conductance) but does not indicate what has been quantified. Similarly, Micu and Plummer (2010) presented results for "skin conductance" (p. 148) without providing any further information on what they quantified. Kim and Fesenmaier (2015) indicated analyzing SCL and SCR but reported "mean EDA value" (p. 423) in the analysis, which suggests they aggregated both tonic EDA (i.e., SCL) value and phasic EDA (i.e., SCR) value. Fox et al. (2018), by referring to both "amplitude of GSR signal heights" and "amplitudes of SC" (p. 49),

were unclear about what exact EDA metric they opted for. Finally, Guerreiro et al. (2015, p. 1740) reported computing the "average skin conductance response level" and abbreviated it as "SCL Mean." This can be confusing because SCL is the standardized abbreviation for skin conductance *level*, not for skin conductance response (Boucsein, 2012). We strongly encourage researchers to use abbreviations proposed by a terminology commission of the Society of Psychophysiological Research (in Boucsein, 2012) to facilitate the communication of their results.

Once computed, these metrics can be subject to statistical analysis. However, they raise concerns about some statistical tests. A first concern relates to inter-individual differences in SCL and in SCR amplitude. An absolute SCR amplitude, per se, offers little insight because a given SCR amplitude can correspond to a large response for one person and to a small response for another person (Boucsein et al., 2012). Therefore, it is recommended to transform the EDA metrics to make them more comparable across individuals (Dawson et al., 2007).

Various strategies to account for these inter-individual differences have been documented in consumer studies. To make EDA measurements comparable across individuals, Guerreiro et al. (2015) subtracted a baseline (pre-experimental, measured during a resting period) from the skin conductance value during the experiment, and cited Lykken and Venables (1971) to justify this transformation. Adam et al. (2015, p. 474) reported using the approach used by Astor, Adam, Jähnig, and Seifert (2013), which consists of dividing the SCR amplitudes by the average SCR amplitude in response to a "neutral reference event" – in their study, the notification that an auction starts. Vanden Abeele and MacLachlan (1994), Lajante et al. (2012), Bettiga et al. (2017), and Shoval et al. (2018) standardized their EDA metrics (i.e., computed the z-score), as suggested by Ben-Shakhar, Lieblich, and Kugelmass (1975) and Ben-Shakhar (1985). Hurley et al. (2015) performed the range-correction transformation proposed by Lykken, Rose, Luther, and

Maley (1966). While consumer studies report multiple ways to account for inter-individual differences in skin conductance values, the psychophysiology literature (e.g., Boucsein et al., 2012; Dawson et al., 2007) refers to standardization and to range-correction transformation only. We thus recommend that consumer researchers use one or the other.

A second concern for the statistical analysis of EDA data relates to the distribution of SCR amplitudes, which is often positively skewed (Benedek & Kaernbach, 2010a, 2010b; Boucsein, 2012). Typically in an EDA recording, most SCRs have a low amplitude and few SCRs have a high amplitude, which results in a positively-skewed distribution of the SCR amplitudes. Therefore, some transformations might be performed to obtain a non-skewed distribution if that is a requirement for the statistical analysis. Venables and Christie (1980 as cited in Boucsein, 2012, p. 178) proposed a log transformation, which was used in multiple consumer studies (e.g., Adam et al., 2015; Bettiga et al., 2017; Droulers et al., 2013; Lajante et al., 2012; Maxian et al., 2013). A square-root transformation is a possible alternative to reduce the positive skewness of the distribution of SCR amplitudes (Boucsein et al., 2012; Dawson et al., 2007).

5.5. Call for transparency

There are multiple ways to record and to analyze EDA data. Some practices are more common than others (e.g., exosomatic recording is widely used, but endosomatic recording is not), some transformations might be particularly recommended (e.g., artifact removal), but there is no gold standard for recording and analyzing EDA data. Each study is unique and researchers need to adapt their recording and analytical practices to the study context. To determine what practice is best suited to the context, researchers should consult the literature in psychophysiology (e.g., Boucsein, 2012).

We would like to stress the importance of reporting all relevant information regarding the recording and analysis of EDA data (Boucsein et al., 2012). Given that multiple EDA recording and analytical practices exist, it is important to report the chosen practice so that readers can assess the validity of the findings and, possibly, replicate them. Table 3 lists the major points to cover when reporting EDA measurements in a study.

--- Table 3 about here ----

Researchers must ensure to detail how they proceeded from recording EDA data to computing EDA metrics. Thoroughly reporting the methods chosen for the EDA study is critical, because missing information prevents the reader from accurately assessing the results of the study. For instance, Hurley et al. (2015) reported that their EDA data analysis detected the occurrence of SCRs (in the time window corresponding to stimulus viewing) for only 3 of 36 participants. The absence of detected SCR for most participants led them to conclude that EDA measurement is not well-suited for capturing the emotional reaction to packaging – the stimulus in their study. However, the authors did not report the criteria they set to detect SCRs. They mentioned only the use of Affectiva's Q-Analytics software. Knowing which minimum amplitude (if any) was chosen to perform the SCR detection could help readers assess their conclusion's validity. The absence of SCRs might not be due to the EDA measurement per se, but rather to the choice of the minimum amplitude. Given that the criteria set to detect SCRs were not reported, this alternative explanation cannot be ruled out.

The example above shows how important it is to thoroughly report the decisions the researchers made to record and analyze EDA data. As many as 23 of the 27 studies (85%) listed in Table 2 did not report one or several fundamental aspects of their studies (type of EDA measurement, preprocessing of EDA data, software used, and/or method to detect and quantify

phasic EDA). Therefore, we call for more transparency when reporting the results of a study with EDA data, for the purposes of validity assessment and replication.

6. Conclusion

EDA measurement offers a great opportunity to overcome the limitations inherent to self-reports of emotions. Since EDA is a psychophysiological measure, it captures the emotional response at the root, removing verbalization and recall biases. We particularly recommend that EDA measurement be used in the following cases: when the study aims to capture how emotional arousal unfolds over time in a certain situation, when social desirability bias is likely to impact self-reports of emotions, and when respondents are unlikely to be able to accurately verbalize and/or recall their emotions. More broadly, EDA measurement can be used to complement self-reports in order to obtain a more complete picture of consumer emotions, a picture that comprises both the physiological and subjective aspects of emotions.

To correctly use EDA measurement, consumer researchers must be aware of the following challenges. EDA data do not indicate the valence of the emotion; they indicate only the degree of emotional arousal. Additional measurements are thus necessary to identify the valence of the emotion being experienced. Moreover, a 1-to-4-second window separates the exposure to a stimulus and the change in EDA level it generates. This window complicates the identification of the stimulus that generated a response when participants were exposed to multiple stimuli within a short time interval. To avoid such a problem, if a study investigates the emotional responses to various stimuli, each stimulus should be administered to participants at spaced-out time intervals. Furthermore, studies that record EDA must be designed to control for attention and cognitive processing because EDA captures these phenomena in addition to emotional arousal.

Importantly, the use of EDA measurement requires specialized equipment to record EDA and specialized procedures to analyze data – procedures that take into account the physiological

properties of EDA. This review shows there are multiple ways to proceed with recording and analyzing EDA data. Therefore, using EDA measurement requires making a series of decisions on how to record, process, analyze, and transform EDA data. To assist marketing researchers in deciding how to record, process, analyze, and transform EDA data, Table 4 summarizes the recommendations that we provided throughout the paper.

--- Table 4 about here ----

There are two major pitfalls in consumer studies using EDA measurement. The first pitfall is to record and analyze EDA data without following procedures established in prior research on EDA. When employing EDA measurement, researchers should consult the literature on psychophysiology (e.g., Boucsein, 2012; Boucsein et al., 2012) to gain relevant insights on this measurement method. The second pitfall is to not communicate effectively or to simply fail to report the procedural and analytical decisions made. Researchers should detail how they proceeded to collect and analyze EDA data so as not to leave room for interpretation.

This paper provides managers and scholars with directions for using EDA measurement in studies of consumers' emotional arousal. This contribution is particularly important given the growing need for new types of insights, including physiological ones, in marketing and consumer research (De Keyser et al., 2015; Lemon & Verhoef, 2016; Morales et al., 2017). Our discussion of EDA measurement will help researchers make an informed decision on whether and how to include it in their studies. Also, this review will assist researchers in turning EDA data into meaningful insights.

Appendix: The psychological significance of EDA

Given that emotionally arousing stimuli result in changes in skin conductance value – via the activation of eccrine sweat glands, EDA measurement informs about the emotional state (Bradley & Lang, 1999). The quantity of sweat produced by the eccrine sweat glands reflects the degree of activation (Stern et al., 2000). Quantifying the degree of arousal generated by a stimulus thus consists of measuring the quantity of sweat that has been secreted, as indicated by the rise in the skin conductance value (Bradley & Lang, 1999; Kroeber-Riel, 1979; Wang & Minor, 2008).

Bradley and Lang (1999) and Sequeira et al. (2009) summarize empirical evidence showing a positive correlation between EDA values and self-reported arousal, with both tonic and phasic EDA measured. Notably, Lang, Greenwald, Bradley, and Hamm (1993) showed that the arousal rating of pictures (from the International Affective Picture System) positively correlates with SCR amplitudes, thus validating the fact that phasic EDA indicates the emotional arousal level. In addition, changes in skin conductance have been shown to significantly correlate with pupil dilation and gastric myoelectrical activity, two other physiological indicators of emotional arousal (Bradley, Miccoli, Escrig, & Lang, 2008; Vianna & Tranel, 2006). As Critchley (2002, p. 133) puts it, "EDA is a widely used and sensitive index of emotion-related sympathetic activity."

Note that eccrine sweat glands are innervated by the sympathetic nervous system, but not by the parasympathetic nervous system (Bradley & Lang, 1999; Critchley, 2002; Larkin, 2006; Stern et al., 2000). Given the different roles that the sympathetic and parasympathetic nervous systems play – the former is responsible for the fight or flight response whereas the latter controls the rest and digest equilibrium – this sole innervation by the sympathetic nervous system means that changes in EDA unambiguously relate to arousal (Norman, Necka, & Berntson, 2016).

References (Appendix)

Bradley, M. M., & Lang, P. J. (1999). Measuring emotion: Behavior, feeling, and physiology. *Cognitive Neuroscience of Emotion*, 246–276.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607.

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *The Neuroscientist*, *8*(2), 132–142.

Kroeber-Riel, W. (1979). Activation research: Psychobiological approaches in consumer research. *Journal of Consumer Research*, *5*(4), 240–250.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273.

Larkin, K. (2006). Psychophysiological assessment. In M. Hersen (Ed.), *Clinician's handbook of adult behavioral assessment* (pp. 165–186). Boston, MA: Elsevier Academic Press.

Norman, G. J., Necka, E., & Berntson, G. G. (2016). The psychophysiology of emotions. In H. Meiselman (Ed.), *Emotion measurement* (pp. 83–98). Cambridge, UK: Woodhead Publishing.

Sequeira, H., Hot, P., Silvert, L., & Delplanque, S. (2009). Electrical autonomic correlates of emotion. *International Journal of Psychophysiology*, *71*(1), 50–56.

Stern, R. M., Ray, W. J., & Quigley, K. S. (2000). *Psychophysiological Recording.* New York, NY: Oxford University Press.

Vianna, E. P., & Tranel, D. (2006). Gastric myoelectrical activity as an index of emotional arousal. *International Journal of Psychophysiology*, *61*(1), 70–76.

Wang, Y. J., & Minor, M. S. (2008). Validity, reliability, and applicability of psychophysiological techniques in marketing research. *Psychology & Marketing*, *25*(2), 197–232.

References

Aaker, D. A., Stayman, D. M., & Hagerty, M. R. (1986). Warmth in advertising: Measurement, impact, and sequence effects. *Journal of Consumer Research*, *12*(4), 365–381.

Abbott, L. (1955). *Quality and competition*. New York, NY: Columbia University Press.

Adam, M. T., Krämer, J., & Müller, M. B. (2015). Auction fever! How time pressure and social competition affect bidders' arousal and bids in retail auctions. *Journal of Retailing*, *91*(3), 468–485.

Alba, J. W., & Williams, E. F. (2013). Pleasure principles: A review of research on hedonic consumption. *Journal of Consumer Psychology*, *23*(1), 2–18.

Alexander, D. M., Trengove, C., Johnston, P., Cooper, T., August, J. P., & Gordon, E. (2005). Separating individual skin conductance responses in a short interstimulus-interval paradigm. *Journal of Neuroscience Methods*, *146*(1), 116–123.

Alexander, V., Tripp, S., & Zak, P. J. (2015). Preliminary evidence for the neurophysiologic effects of online coupons: Changes in oxytocin, stress, and mood. *Psychology & Marketing*, *32*(9), 977–986.

Astor, P. J., Adam, M. T., Jähnig, C., & Seifert, S. (2013). The joy of winning and the frustration of losing: A psychophysiological analysis of emotions in first-price sealed-bid auctions. *Journal of Neuroscience, Psychology, and Economics*, *6*(1), 14.

Bach, D. R., Daunizeau, J., Friston, K. J., & Dolan, R. J. (2010). Dynamic causal modelling of anticipatory skin conductance responses. *Biological psychology*, *85*(1), 163-170.

Bach, D. R., Friston, K. J., & Dolan, R. J. (2013). An improved algorithm for model-based analysis of evoked skin conductance responses. *Biological psychology*, *94*(3), 490-497.

Baumgartner, H., Sujan, M., & Padgett, D. (1997). Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research*, 219–232.

Ben-Shakhar, G. (1985). Standardization within individuals: A simple method to neutralize individual differences in skin conductance. *Psychophysiology*, *22*(3), 292–299.

Ben-Shakhar, G., Lieblich, I., & Kugelmass, S. (1975). Detection of information and GSR habituation: An attempt to derive detection efficiency from two habituation curves. *Psychophysiology*, *12*(3), 283–288.

Benedek, M. (n.d.). Analysis of EDA data using Ledalab [PDF]. Retrieved from http://www.ledalab.de/download/Analysis%20of%20EDA%20data%20using%20Ledalab.pdf

Benedek, M., & Kaernbach, C. (2010a). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91.

Benedek, M., & Kaernbach, C. (2010b). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, *47*(4), 647–658.

Bettiga, D., Lamberti, L., & Noci, G. (2017). Do mind and body agree? Unconscious versus conscious arousal in product attitude formation. *Journal of Business Research*, *75*, 108–117.

Bolls, P. D., Lang, A., & Potter, R. F. (2001). The effects of message valence and listener arousal on attention, memory, and facial muscular responses to radio advertisements. *Communication Research*, *28*(5), 627–651.

Bolls, P. D., Muehling, D. D., & Yoon, K. (2003). The effects of television commercial pacing on viewers' attention and memory. *Journal of Marketing Communications*, *9*(1), 17–28.

Boshoff, C. (2012). A neurophysiological assessment of consumers' emotional responses to service recovery behaviors: The impact of ethnic group and gender similarity. *Journal of Service Research*, *15*(4), 401–413.

Boshoff, C. (2017). An assessment of consumers' subconscious responses to frontline employees' attractiveness in a service failure and recovery situation. *South African Journal of Economic and Management Sciences*, *20*(1), 1–13.

Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034.

Bradley, M. M., & Lang, P. J. (1999). Measuring emotion: Behavior, feeling, and physiology. *Cognitive Neuroscience of Emotion*, 246–276.

Bradley, S. D., Angelini, J. R., & Lee, S. (2007). Psychophysiological and memory effects of negative political ads: Aversive, arousing, and well remembered. *Journal of Advertising*, *36*(4), 115–127.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 173–191). New York, NY: Guilford Press.

Cowley, E. (2008). Looking back at an experience through rose-colored glasses. *Journal of Business Research*, *61*(10), 1046–1052.

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *The Neuroscientist*, *8*(2), 132–142.

Dalakas, V. (2006). The importance of a good ending in a service encounter. *Services Marketing Quarterly*, *28*(1), 35–53.

Damasio, A. R. (2000). A second chance for emotion. *Cognitive Neuroscience of Emotion*, 12–23.

Dawson, M. E., Schell, A. M., & Courtney, C. G. (2011). The skin conductance response, anticipation, and decision-making. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 111.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159–181). New York, NY: Cambridge University Press.

De Chernatony, L. (2010). *From brand vision to brand evaluation: The strategic process of growing and strengthening brands*. Routledge.

De Keyser, A., Lemon, K. N., Klaus, P., & Keiningham, T. L. (2015). A framework for understanding and managing the customer experience. *Marketing Science Institute Working Paper Series*, 15–121.

Droulers, O., Lajante, M., & Lacoste-Badie, S. (2013). Apport de la démarche neuroscientifique à la mesure des émotions: Importation d'une nouvelle méthode de mesure de l'activité électrodermale. *Décisions Marketing*, (72), 87.

Fox, A. K., Deitz, G. D., Royne, M. B., & Fox, J. D. (2018). The face of contagion: Consumer response to service failure depiction in online reviews. *European Journal of Marketing*, *52*(1/2), 39–65.

Fredrickson, B. L., & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition & Emotion*, *12*(2), 191–220.

Gakhal, B., & Senior, C. (2008). Examining the influence of fame in the presence of beauty: An electrodermal "neuromarketing" study. *Journal of Consumer Behaviour*, *7*(4–5), 331–341.

Galak, J., & Meyvis, T. (2011). The pain was greater if it will happen again: The effect of anticipated continuation on retrospective discomfort. *Journal of Experimental Psychology: General*, *140*(1), 63–75.

Greco, A., Valenza, G., Lanata, A., Scilingo, E. P., & Citi, L. (2015). cvxEDA: A convex optimization approach to electrodermal activity processing. *IEEE Transactions on Biomedical Engineering*, *63*(4), 797-804.

Groeppel-Klein, A. (2005). Arousal and consumer in-store behavior. *Brain Research Bulletin*, *67*(5), 428–437.

Groeppel-Klein, A., & Baun, D. (2001). The role of customers' arousal for retail store – results from an experimental pilot study using electrodermal activity as indicator. *ACR North American Advances*.

Gross, J. J. (1998). Antecedent-and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224.

Guerreiro, J., Rita, P., & Trigueiros, D. (2015). Attention, emotions and cause-related marketing effectiveness. *European Journal of Marketing*, *49*(11/12), 1728–1750.

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic consumption: Emerging concepts, methods and propositions. *The Journal of Marketing*, 92–101.

Hurley, R. A., Hutcherson, D. E., Tonkin, C. E., Dailey, S. B., & Rice, J. C. (2015). Measuring physiological arousal towards packaging: Tracking electrodermal activity within the consumer shopping environment. *Journal of Applied Packaging Research*, *7*(3), 5.

Kilbourne, W. E., Painton, S., & Ridley, D. (1985). The effect of sexual embedding on responses to magazine advertisements. *Journal of Advertising*, *14*(2), 48–56.

Kim, J., & Fesenmaier, D. R. (2015). Measuring emotions in real time: Implications for tourism experience design. *Journal of Travel Research*, *54*(4), 419–429.

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, *74*(3), 686.

Kroeber-Riel, W. (1979). Activation research: Psychobiological approaches in consumer research. *Journal of Consumer Research*, *5*(4), 240–250.

LaBarbera, P. A., & Tucciarone, J. D. (1995). GSR reconsidered: A behavior-based approach to evaluating and improving the sales potency of advertising. *Journal of Advertising Research*, *35*(5), 33–54.

Lajante, M., Droulers, O., Dondaine, T., & Amarantini, D. (2012). Opening the "black box" of electrodermal activity in consumer neuroscience research. *Journal of Neuroscience, Psychology, and Economics*, *5*(4), 238.

Lang, P. J. (1993). The three-system approach to emotion. In N. Birbaumer & A. Öhman (Eds.), *The structure of emotion: Psychophysiological, cognitive, and clinical aspects* (pp. 18-30). Seattle, WA: Hogrefe & Huber.

Langner, T., Schmidt, J., & Fischer, A. (2015). Is it really love? A comparative investigation of the emotional nature of brand and interpersonal love. *Psychology & Marketing*, *32*(6), 624–634.

Larkin, K. (2006). Psychophysiological Assessment. In M. Hersen (Ed.), *Clinician's handbook of adult behavioral assessment* (pp. 165–186). Boston, MA: Elsevier Academic Press.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, *46*(8), 819.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*(6), 69–96.

Lykken, D. T., Rose, R., Luther, B., & Maley, M. (1966). Correcting psychophysiological measures for individual differences in range. *Psychological Bulletin*, *66*(6), 481.

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, *8*(5), 656–672.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237.

Maxian, W., Bradley, S. D., Wise, W., & Toulouse, E. N. (2013). Brand love is in the heart: Physiological responding to advertised brands. *Psychology & Marketing*, *30*(6), 469–478.

Micu, A. C., & Plummer, J. T. (2010). Measurable emotions: How television ads really work: Patterns of reactions to commercials can demonstrate advertising effectiveness. *Journal of Advertising Research*, *50*(2), 137–153.

Morales, A. C., Amir, O., & Lee, L. (2017). Keeping It Real in Experimental Research—Understanding When, Where, and How to Enhance Realism and Measure Consumer Behavior. *Journal of Consumer Research*.

Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2009). Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. *Journal of Neuroscience, Psychology, and Economics*, *2*(1), 21.

Picard, R. W., Fedor, S., & Ayzenberg, Y. (2016). Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, *8*(1), 62–75.

Plassmann, H., Venkatraman, V., Huettel, S., & Yoon, C. (2015). Consumer neuroscience: applications, challenges, and possible solutions. *Journal of Marketing Research*, *52*(4), 427–435.

Poels, K., & Dewitte, S. (2006). How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research*, *46*(1), 18–37.

Reimann, M., Castaño, R., Zaichkowsky, J., & Bechara, A. (2012). How we relate to brands: Psychological and neurophysiological insights into consumer–brand relationships. *Journal of Consumer Psychology*, *22*(1), 128–142.

Rienecker, L., Jørgensen, P. S., & Skov, S. (2013). *The good paper: A handbook for writing papers in higher education*. International Specialized Book Service Incorporated.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Sequeira, H., Hot, P., Silvert, L., & Delplanque, S. (2009). Electrical autonomic correlates of emotion. *International Journal of Psychophysiology*, *71*(1), 50–56.

Shoval, N., Schvimer, Y., & Tamir, M. (2018). Real-time measurement of tourists' objective and subjective emotions in time and space. *Journal of Travel Research*, *57*(1), 3–16.

Somervuori, O., & Ravaja, N. (2013). Purchase behavior and psychophysiological responses to different price levels. *Psychology & Marketing*, *30*(6), 479–489.

Stern, R. M., Ray, W. J., & Quigley, K. S. (2000). *Psychophysiological Recording.* New York, NY: Oxford University Press.

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015, August). Automatic identification of artifacts in electrodermal activity data. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 1934–1937). IEEE.

Topolinski, S., Lindner, S., & Freudenberg, A. (2014). Popcorn in the cinema: Oral interference sabotages advertising effects. *Journal of Consumer Psychology*, *24*(2), 169–176.

van Dooren, M., de Vries, J.J.G., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior*, *106*(2), 298–304.

Vanden Abeele, P., & MacLachlan, D. L. (1994). Process tracing of emotional responses to TV ads: Revisiting the warmth monitor. *Journal of Consumer Research*, *20*(4), 586–600.

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., & Winer, R. S. (2015). Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling. *Journal of Marketing Research*, *52*(4), 436–452.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, *85*(1), 31–41.

Walla, P., Brenner, G., & Koller, M. (2011). Objective measures of emotion related to brand attitude: A new way to quantify emotion-related aspects relevant to marketing. *PLOS ONE 6*(11).

Walla, P., Koller, M., Brenner, G., & Bosshard, S. (2017). Evaluative conditioning of established brands: Implicit measures reveal other effects than explicit measures. *Journal of Neuroscience, Psychology, and Economics*, *10*(1), 24.

Wiens, S. (2005). Interoception in emotional experience. *Current Opinion in Neurology*, *18*(4), 442–447.

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science*, *13*(3), 120–123.

Xia, V., Jaques, N., Taylor, S., Fedor, S., & Picard, R. (2015). Active learning for electrodermal activity classification. *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 1–6. https://doi.org/10.1109/SPMB.2015.7405467

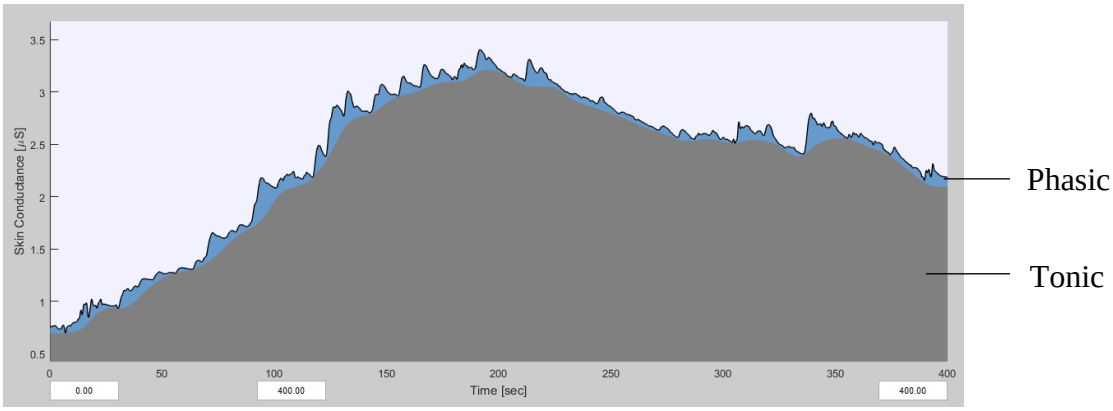

Figure 1. Decomposition of an EDA signal into its tonic and phasic components

a)

b)

c)

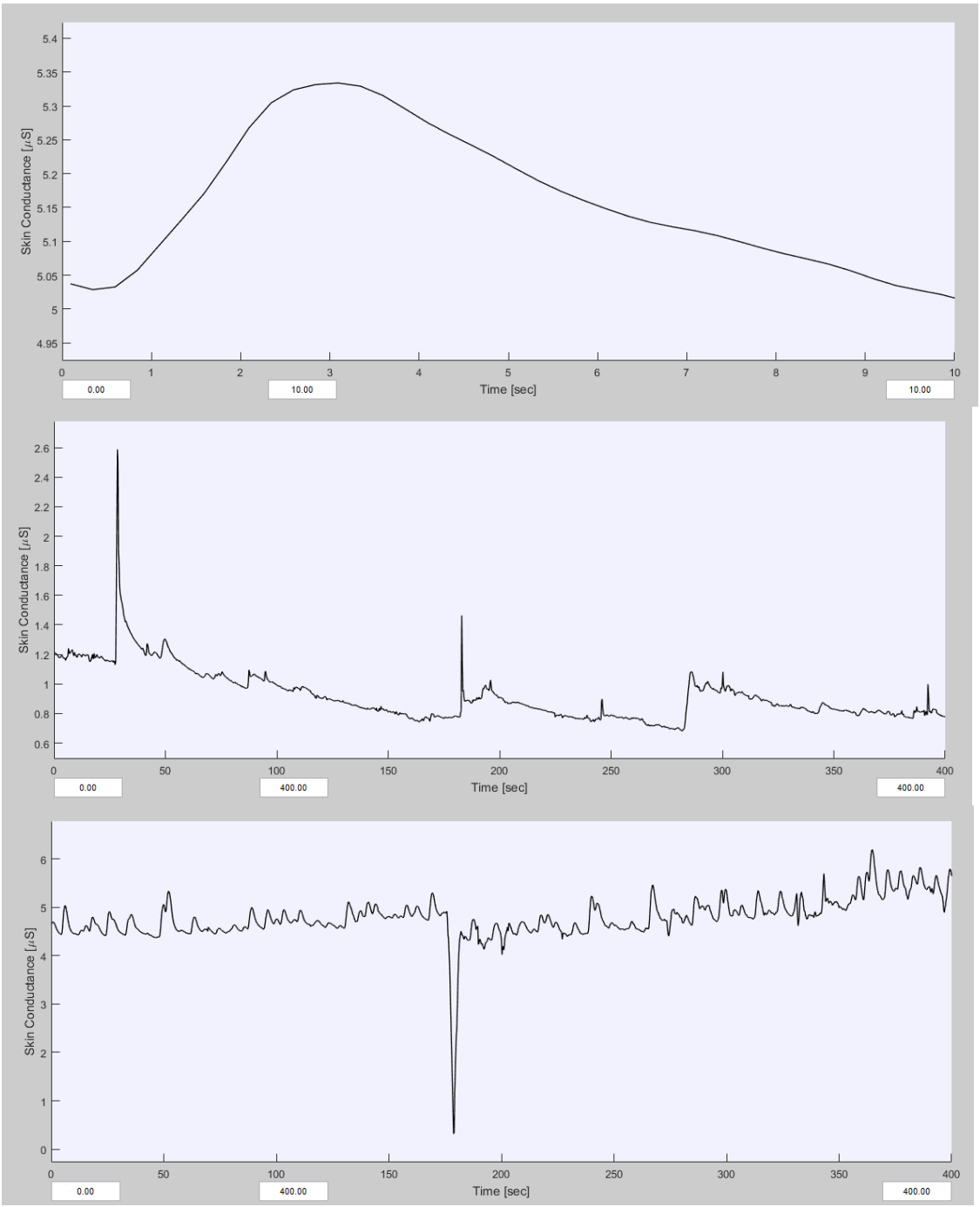

Figure 2. (a) Close-up of the shape of an SCR; (b) EDA signal with movement artifacts; (c) A sudden drop in the EDA signal caused by a loose connection.

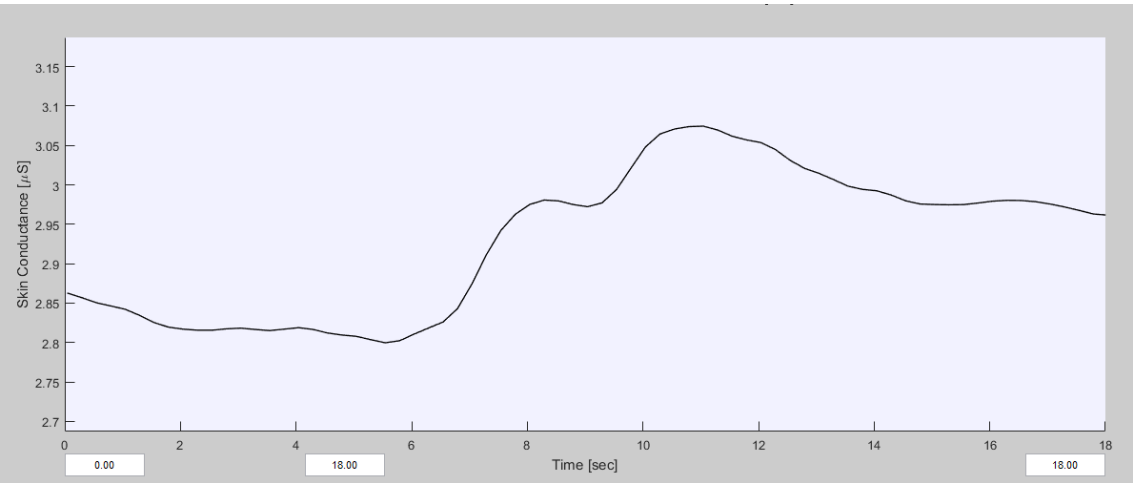

Figure 3. Overlapping SCRs

Table 1. Summary of prior studies of consumer emotions that reported using EDA measurement

Study	Context	Number of participants with EDA measurement	Key findings derived from EDA measurement
Aaker, Stayman, and Hagerty (1986) – Study 1	Advertising	30	Perceived warmth in a TV commercial positively correlates with arousal.
Adam, Krämer, and Müller (2015) – Study 1	Retail auctions	240	Winning an ascending clock auction is more arousing than losing it.
Bettiga, Lamberti, and Noci (2017)	Attitude toward a product	160	Arousal elicited by the interaction with a product has a negative effect on attitude toward this product when the interaction is direct; it has no effect on attitude when the interaction is virtual.
Bolls, Lang, and Potter (2001)	Advertising	41	Positive messages in radio advertisements are more arousing than negative messages.
Boshoff (2012)	Service recovery	64	Service recovery performed by a service employee of the same or opposite gender, or of the same or a different ethnicity is not arousing.
Boshoff (2017)	Service recovery	40	Service recovery performed by a physically attractive service employee of the opposite gender is less arousing than service recovery performed by a less attractive employee.
Droulers, Lajante, and Lacoste-Badie (2013)	Advertising	48	Some TV commercials are more arousing than others.
Fox, Deitz, Royne, and Fox (2018) – Study 1	Online consumer reviews	56	Reading a negative consumer review is more arousing than reading a positive or neutral consumer review.
Gakhal and Senior (2008)	Advertising	24	Ads showing a celebrity are more arousing than ads showing a non-celebrity. This effect of fame on arousal is stronger when the celebrity is average-looking rather than attractive.
Groeppel-Klein (2005)	Shopping environments	27 (Study 1*); Not reported (Study 2); 82 (Study 3); 105 (Study 4)	A store or display in which merchandising principles have been applied is more arousing than a store or display in which these principles have not been applied (Studies 1 and 2). Buyers in a shopping environment are more aroused than non-buyers (Studies 3 and 4).
Groeppel-Klein and Baun (2001)	Shopping environments	27	A store in which environmental psychology principles have been applied is more arousing than a store in which such principles have not been applied.
Guerreiro, Rita, and Trigueiros (2015)	Purchase decisions	48	Arousal elicited by a cause-related product predicts the choice of this product if this product is hedonic but not if it is utilitarian.
Hurley, Hutcherson, Tonkin, Dailey, and Rice (2015)	Packaging	42	Viewing packaging is not arousing.

Kim and Fesenmaier (2015)	Touristic activities	2	Different touristic activities elicit different levels of arousal.
Kroeber-Riel (1979)	Advertising	60 (study by Witt); 84 (study by Barg)	Arousal elicited by ads increases the pick-up of information in these ads (study by Witt); and it increases the recall of information contained in the ads (study by Barg).
Lajante, Droulers, Dondaine, and Amarantini (2012)	Advertising	30	Some TV commercials are more arousing than others.
Langner, Schmidt, and Fischer (2015)	Brands	7	A loved person is more arousing than a loved brand. A liked person is not more arousing than a loved brand. A loved brand is more arousing than a liked brand.
Maxian, Bradley, Wise, and Toulouse (2013)	Brands	56	More-loved brands are not more arousing than less-loved brands.
Micu and Plummer (2010)	Advertising	50	An argument-based TV commercial is more arousing than an emotional-based one.
Ohme, Reykowska, Wiener, and Choromanska (2009)	Advertising	45	A TV commercial scene in which a female model gestures is more arousing than the same scene lacking a gesture.
Reimann, Castaño, Zaichkowsky, and Bechara (2012) – Study 2	Brands	25	Brands with which consumers have recently formed a close relationship are more arousing than are brands with which consumers long ago formed a close relationship.
Shoval, Schvimer, and Tamir (2018)	Touristic activities	68	Different touristic sites elicit different levels of arousal.
Somervuori and Ravaja (2013)	Purchase decisions	33	Arousal elicited by viewing a product does not predict the decision to purchase this product.
Vanden Abeele and MacLachlan (1994)	Advertising	51	Perceived warmth in a TV commercial does not correlate with arousal.
Venkatraman et al. (2015)	Advertising	29	Arousal elicited by TV commercials does not predict the market-level advertising elasticities.
Walla, Brenner, and Koller (2011)	Brands	29	Disliked brands are more arousing than liked brands.
Walla, Koller, Brenner, and Bosshard (2017)	Brands	21	Evaluative conditioning of brands does not change the level of arousal elicited by these brands.

*Study 1 in Groeppel-Klein (2005) corresponds to the study reported in Groeppel-Klein and Baun (2001).

Table 2. Recording and analytical practices in studies of consumer emotions that employ EDA measurement

Study	Type of EDA measurement	Preprocessing of EDA data	Software used to record and/or process EDA data	Method to detect and quantify phasic EDA	EDA metrics
Aaker, Stayman, and Hagerty (1986) – Study 1	Skin resistance	Not reported	Not applicable	Not reported	"Resistance amplitudes" (p. 369)
Adam, Krämer, and Müller (2015) – Study 1	Skin conductance	Not reported	Ledalab	Not reported	SCR amplitude
Bettiga, Lamberti, and Noci (2017)	Skin conductance	Artifact removal	Ledalab	Continuous decomposition analysis	Sum of SCR amplitudes; average value of the phasic response; and ISCR
Bolls, Lang, and Potter (2001)	Skin conductance	Not reported	Not reported	Not reported	Number of SCRs
Boshoff (2012)	Skin conductance	Downsampling; filtering	Not reported	"Differential analysis, wavelet transformation, and other mathematical and statistical tools" (p. 404)	"Galvanic skin response" (p. 406)
Boshoff (2017)	Not reported	Not reported	Not reported	Not reported	"GSR index" (p. 7)
Droulers, Lajante, and Lacoste-Badie (2013)	Skin conductance	Smoothing; artifact removal	Ledalab V3.3.2	Continuous decomposition analysis	ISCR
Fox, Deitz, Royne, and Fox (2018)	Skin conductance	Filtering	Not reported	Not reported	"Amplitudes of GSR signal heights" (p. 49)
Gakhal and Senior (2008)	Skin conductance	Not reported	Not reported	Baseline	"EDA per cent signal change" (p. 335)
Groeppel-Klein (2005)	Skin conductance	Artifact exclusion	Not reported	Not applicable (EDR registered during recording)	Sum of SCR amplitudes; frequency of SCRs
Groeppel-Klein and Baun (2001)	Skin conductance	Artifact exclusion	Not reported	Not reported	Sum of SCR amplitudes; sum of SCR amplitudes/minute; frequency of SCRs
Guerreiro, Rita, and Trigueiros (2015)	Skin conductance	Filtering	Not reported	Not reported	"Average skin conductance response level"; "average time to reach the emotional response peak" (p. 1740)
Hurley, Hutcherson, Tonkin, Dailey, and Rice (2015)	Skin conductance	Smoothing	Affectiva Q-Analytics	Not reported	Presence/absence of SCRs; mean EDA value

Kim and Fesenmaier (2015)	Skin conductance	Artifact correction	MATLAB; SPSS 19.0	Not reported	"Mean EDA value" (p. 423)
Kroeber-Riel (1979)	Skin resistance	Not reported	Not applicable	Not reported	EDR amplitude
Lajante, Droulers, Dondaine, and Amarantini (2012)	Skin conductance	Downsampling; smoothing; artifact removal	Ledalab V3.3.2	Continuous decomposition analysis	ISCR; sum of SCR amplitudes
Langner, Schmidt, and Fischer (2015)	Skin conductance	Filtering	Acqknowledge	Acqknowledge's EDA data analysis	SCR amplitude; SCR rise time; SCR recovery time
Maxian, Bradley, Wise, and Toulouse (2013)	Skin conductance	Not reported	Not reported	Baseline	"Maximum change" (p. 474)
Micu and Plummer (2010)	Skin conductance	Not reported	Not reported	Not reported	"Skin conductance" (p. 148)
Ohme, Reykowska, Wiener, and Choromanska (2009)	Skin conductance	Downsampling; filtering	Not reported	"Differential analysis, wavelet transformation, and other mathematical and statistical tools" (p. 26)	"Averaged skin conductance response" (p. 28)
Reimann, Castaño, Zaichkowsky, and Bechara (2012) – Study 2	Skin conductance	Not reported	Not reported	Baseline	"Normalized SCR" (p. 134)
Shoval, Schvimer, and Tamir (2018)	Skin conductance	Not reported	Not reported	Not relevant (tonic EDA is of interest)	SCL values
Somervuori and Ravaja (2013)	Skin conductance	Not reported	Psylab7	Baseline	"EDA delta scores" (p. 484)
Vanden Abeele and MacLachlan (1994)	Skin conductance	Not reported	INTERTEST; ROGIL	Not reported	"GSR score" (p. 592)
Venkatraman et al. (2015)	Skin conductance	Not reported	Acqknowledge 4.0	Acqknowledge's built-in EDA exploratory data analysis (high-pass filtering)	SCR amplitude
Walla, Brenner, and Koller (2011)	Skin conductance	Not reported	Biotrace+	Not reported	"Mean skin conductance values" (p. 3)
Walla, Koller, Brenner, and Bosshard (2017)	Skin conductance	Not reported	Biotrace+	Not reported	"Mean skin conductance values" (p. 29)

Table 3. Information to report in a study that employs EDA measurement

Stage	**What to report**	
Recording	Recording conditions	Was EDA recorded in field or lab settings? Were participants moving or still?
	Recording technique	Was the EDA recording endosomatic or exosomatic? If exosomatic recording, was direct or alternating current applied? What electrical measure (e.g., conductance, resistance) was recorded? Was an AC-coupled amplifier used to record the phasic component of EDA alone?
	Electrode placement	On what part of the body were the electrodes placed?
Preprocessing	Smoothing/Low-pass filtering	Was any smoothing or low-pass filter applied to the raw EDA signal?
	Artifact detection and removal	How were artifacts detected (visual or automated inspection)? Were data points corresponding to the artifacts removed and reconstructed or discarded?
Detection and quantification of phasic EDA	Analytical method	If no AC-coupled amplifier was used during recording, what analytical method (e.g., trough-to-peak analysis, continuous decomposition analysis) was used to detect phasic EDA? If an AC-coupled amplifier separated phasic from tonic EDA during the recording, please explicitly state it.
	Criteria	What minimum amplitude for SCRs was set for detecting and quantifying SCRs? Was any other criterion set?
	Time window	Were SCRs detected during a specific time window (e.g., 1-to-4 seconds after stimulus onset) or during an extended period of time (e.g., entire recording)?
	Possible issues (overlapping SCRs, habituation)	Were overlapping SCRs spotted during visual inspection of the EDA signal? What solution was used to detect and quantify overlapping SCRs? If the study employed a repeated-measure design, was habituation of the SCRs taken into consideration?
Computation of EDA metrics	Choice of the metrics of interest	What quantification of phasic EDA (e.g., SCR amplitude, area under the curve) was used as an indicator of arousal?
	Transformation performed on the metrics	Were the EDA metrics transformed to account for inter-individual differences? If yes, what transformation (e.g., standardization) was performed? Were the EDA metrics transformed (e.g., log-transformation) to reduce the skewness of their distributions?

Table 4. Summary of recommendations for employing EDA measurement

Stage	**Recommendations**	
Recording	Recording conditions	A controlled environment is preferred to a less controlled environment, and participants should preferably stay still during the recording. However, the purpose of the study might require participants to move and/or the recording to occur in a specific environment (e.g. field study).
	Recording technique	What technique is used depends on the EDA recording equipment. Usual equipment performs an exosomatic recording with DC or with AC converted to DC and provides a conductance measure.
	Electrode placement	Palms and fingers of the non-dominant hand are preferred to other locations for electrode placement.
Preprocessing	Smoothing/Low-pass filtering	Smoothing or low-pass filtering (e.g., Butterworth filter) the raw EDA signal is recommended to remove high-frequency noise and small artifacts.
	Artifact detection and removal	A visual or automated inspection of the EDA signal should be performed to detect artifacts. The data points corresponding to detected artifacts need to be corrected or discarded.
Detection and quantification of phasic EDA	Analytical method	If an AC-coupled amplifier was used to separate the phasic from the tonic component of the EDA signal during the recording, only the quantification of phasic EDA needs to be performed. If no AC-coupled amplifier was used, the phasic component of the EDA signal needs to be detected and quantified. Several methods and/or software programs have been developed to this end (e.g, Alexander et al., 2005; Bach, Daunizeau, Friston, & Dolan, 2010; Bach, Friston, & Dolan, 2013; Benedek & Kaernbach, 2010a, b; Greco, Valenza, Lanata, Scilingo, & Citi, 2015)
	Criteria	What criteria need to be set depends on the analytical method. Setting a minimum threshold for SCRs is usually required, else recommended. This threshold should be set between 0.01 and 0.05 μS depending on the recording conditions (the less controlled the environment, the higher the threshold should be).
	Time window	If the response to one discrete stimulus is of interest, it is recommended to detect SCRs within the 1-to-4-second window following stimulus onset. If the responses to multiple stimuli occurring over time are of interest, it is relevant to detect phasic EDA during the entire period of interest.
	Possible issues (overlapping SCRs, habituation)	If overlapping SCRs were detected by visual inspection of the EDA signal, a suitable analytical method must be chosen; examples of such methods are cvxEDA (Greco et al., 2015, continuous decomposition analysis (Benedek & Kaernbach, 2010a), or the method proposed by Alexander et al., 2005. If a repeated-measure design was employed, it is recommended to test whether the SCR amplitude decreased with each recurrence of the stimulus.
Computation of EDA metrics	Choice of the metrics of interest	If one SCR was detected during the chosen time window, the SCR amplitude or, alternatively, the area under the curve can be used to index emotional arousal. If multiple SCRs were detected, average SCR amplitude or, alternatively, ISCR can be used to index arousal level.
	Transformation performed on the metrics	For more meaningful between-subjects comparisons, standardization or range-correction transformation of the EDA metrics can be performed. If the statistical analysis to be performed on the EDA metrics require the data distribution to not be skewed, a log or square root transformation of the EDA metrics can be performed.

Delphine Caruelle is a PhD candidate in Marketing at BI Norwegian Business School. She is currently writing a dissertation on customers' emotional experience during service encounters. In her research, Ms. Caruelle embraces the use of diverse types of data (e.g., survey data, physiological data) collected in multiple ways (e.g., secondary data sources, experiments, field studies). Prior to joining BI Norwegian Business School, Ms. Caruelle worked as Consumer Insights Specialist at Opera Software. She completed a Research Master in Marketing and Strategy from Paris-Dauphine University in 2013.

Anders Gustafsson is a professor of marketing at the Norwegian Business School. He is also Distinguished Professorial Fellow at the University of Manchester's Alliance Manchester Business School (AMBS). Dr. Gustafsson has published articles in journals such as *Journal of Marketing, Journal of Marketing Research, Journal of Service Research, Journal of Business Research, and Industrial Marketing Management*. Recently he received he received the Christopher Lovelock Career Contributions to the Services Discipline Award.

Dr. Poja Shams is an assistant Professor at Karlstad Business School in Sweden. His research has primarily been focused on consumer decision making and visual attention in the retail environment. His research has been awarded by the Gunnar Sundblad Research Foundation and published in several distinguished journals such as *Journal of Business & Retail Management Research, Journal of Business Research and Psychology & Marketing*.

Line Lervik-Olsen is a professor of marketing at BI Norwegian Business School and holds a part time position at the Norwegian School of Economics and the Center for Service Innovation. Her work is published in journals such as the *Journal of Service Research,* the *Journal of Economic Psychology, Journal of Service Management, Managing Service Quality, Journal of Services Marketing, Journal of Service Theory and Practice, PLOS One, Journal of Business Research,* as well as in books.