*Paweł Korpal**
*Katarzyna Jankowiak**

Physiological and self-report measures in emotion studies: Methodological considerations

Abstract: *Investigating human emotions empirically is still considered to be challenging, mostly due to the questionable validity of the results obtained when employing individual types of measures. Among the most frequently used methods to study emotional reactions are self-report, autonomic, neurophysiological, and behavioral measures. Importantly, previous studies on emotional responding have rarely triangulated the aforementioned research methods. In this paper we discuss main methodological considerations related to the use of physiological and self-report measures in emotion studies, based on our previous research on the processing of emotionally-laden narratives in the native and non-native language, where we employed the SUPIN S30 questionnaire as a self-report tool, and galvanic skin response (GSR) as a physiological measure (Jankowiak & Korpal, 2018). The findings revealed a more pronounced reaction to stimuli presented in the native relative to the non-native language, which was however reflected only in GSR patterns. The lack of correlation between GSR and SUPIN scores might have resulted from a number of methodological considerations, such as social desirability bias, sensitive questions, lack of emotional self-awareness, compromised ecological validity, and laboratory anxiety, all of which are thoroughly discussed in the article.*

Keywords: *galvanic skin response, self-report measure, emotional responding, methodology triangulation*

Introduction

Human emotional response is defined as modulated by a personal perceived significance of a specific event, which results in experiential, physiological as well as behavioral reaction to it (Scherer, 1984; Smith & Ellsworth, 1985; Frijda, 1988; Lang, 1988; Lazarus, 1991; Gross, 2007; Larsen & Prizmic-Larsen, 2006; Mauss & Robinson, 2009). Emotional responding is frequently discussed from either a dimensional or discrete perspective. In line with the dimensional approach, emotional reactions are organized around a few fundamental dimensions, such as valence, arousal, and approach-avoidance (Lang, Bradley, & Cuthbert, 1997; Davidson, 1999; Russell & Barrett, 1999; Watson, Wiese, Vaidya, & Tellegen, 1999). While emotional valence is defined as reflecting either positive or negative states, arousal refers to the magnitude of the emotional response (e.g., low arousal – quiet; high arousal – surprised; Mauss & Robinson, 2009), and the approach-avoidance dimension reflects tendencies to either approach or avoid specific stimuli. The discrete

emotions perspective, on the other hand, posits that each emotion experienced relates to a different experiential, physiological as well as behavioral reaction to it (Ekman, 1999; Panksepp, 2007).

Though over the last decades much attention has been devoted to empirically investigating human emotions, measuring participants' emotional state in an experimental setting is still considered to be challenging, mostly due to the questionable validity of the results obtained when employing individual types of measures. Among the most frequently employed methods for studying emotional reactions are self-report, autonomic, neurophysiological, as well as behavioral measures, each one providing different types of information concerning the dimensional and discrete perspective on emotional responding.

In self-report measures, participants are asked to report their currently experienced emotions. When employing self-reports to examine emotional states, researchers usually use either an immediate or delayed self-report response (Robinson & Clore, 2002). While the former involves a self-report of current emotional reactions, the

---
* **Adam Mickiewicz University in Poznań**

Corresponding author: Paweł Korpal, e-mail: pkorpal@wa.amu.edu.pl

latter pertains to self-reports provided at distanced time from the emotion experienced, and is thus considered to be less valid (Mauss & Robinson, 2009). Self-reports can take different forms, the most frequent being the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), the affect rating dial (Gottman & Levenson, 1985), as well as the positive-negative affect meter (Levenson & Gottman, 1983). The PANAS questionnaire consists of adjectives representing negative (50%) and positive (50%) emotions. Participants are instructed to report their emotional state on a 5-point Likert scale, where "1" indicates that a specific emotion is not experienced, whereas "5" indicates a very strongly experienced emotion. Consequently, participants need to assess their emotional state with respect to all of the emotional labels listed in the questionnaire. The major advantage of such a questionnaire is that it is relatively easy to administer, and to fill in by participants. Yet, it has also received some criticism, mostly due to the fact that administering such a questionnaire interrupts the experiment, and the test itself does not provide continuous data of emotions experienced (Ruef & Levenson, 2007). A continuous measure of emotions that provides a time-course of emotional experience might however be obtained by means of employing the affect rating dial (Gottman & Levenson, 1985), which involves a joystick device with an adjustable dial from 0° to 180° and a 9-point scale, on which participants are asked to report their current emotional state (0° = very negative; 90° = neutral; 180° = very positive). The affect rating dial has been further adopted and developed into a positive-negative affect meter (Levenson & Gottman, 1983), which involves a knob attached to a potentiometer controlling 15 colored lights, which are positioned above the computer screen. The colored lights include 7 green lights (representing a positive emotional state), 7 red lights (representing a negative emotional state) and a yellow light at the center (representing a neutral emotional state). Participants are instructed to adjust the knob according to their currently experienced emotions as often as possible. Both the affect rating dial and the positive-negative affective meter allow for measuring rapid changes in emotions experienced by participants. Notwithstanding such an advantage, these methods also suffer from some limitations. Namely, unlike the PANAS questionnaire, they involve a single-valence dimension of emotion with positive and negative poles, as a result of which they limit participants to choosing only one end of a continuum, even though they might experience both positive (e.g., interested) and negative (e.g., sad) emotions at the same time (Ruef & Levenson, 2007). Still, self-report measurements are advantageous as, unlike other methods discussed below, they can be used to measure both emotional arousal and the valence of the presented stimuli.

Other methods employed when examining emotional responding are sensitive to either emotional valence or arousal. To provide insights into whether stimuli are perceived as evoking positive or negative emotions, facial electromyography (EMG) can be adopted, which requires recording electrical potentials from facial muscles by means of placing several electrodes on the face (Mauss & Robinson, 2009). The electrical potentials are usually recorded from two facial muscle groups: corrugator supercilii and zygomatic muscles. While the activity of corrugator supercilii muscles, which are involved in furrowing the eyebrows, increases as the level of pleasantness decreases, the activity of zygomatic muscles, which are related to raising of the corner of the lips, increases with the level of pleasantness (Mauss & Robinson, 2009). Consequently, the facial EMG is frequently employed when studying implicit evaluation processes (Cacioppo, Berntson, Klein, & Poehlmann, 1997; Dimberg, Thunberg, & Elmehed, 2000).

When investigating emotional arousal, on the other hand, much research has employed electroencephalography (EEG) as a measure of electrical brain activity, which reflects biophysical events from populations of neurons (Cohen, 2014). Among several types of analyses that can be performed based on the recorded EEG signal, much research has employed event-related potential (ERP) analyses, which provide an average brain response to an external event, such as the presentation of an external stimulus (Kutas, Van Petten, & Kluender, 2006). Two ERP components are frequently discussed when examining emotional responding. First, the P300, which is elicited between 300–500 ms post stimulus onset, indexes stimulus salience, with more pronounced P300 amplitudes elicited in response to unexpected, improbable, and task-irrelevant stimuli (Duncan-Johnson & Donchin, 1977; Hajcak, MacNamara, & Olvet, 2010). Second, the late positive potential (LPP) is evoked at around 400–600 ms after the presentation of a stimulus, and can be observed even until 1000 ms after stimulus offset (Schupp, Flaisch, Stockburger, & Junghöfer, 2006; Hajcak & Olvet, 2008; Brown, van Steenbergen, Band, de Rover, & Nieuwenhuis, 2012). In studies on emotion processing, both the P300 and the LPP have been found to be modulated by emotional intensity of a stimulus, as emotionally-laden materials (irrespectively of their valence) have been observed to elicit larger P300 and LPP amplitudes compared to neutral conditions (Naumann, Bartussek, Diedrich, & Laufer, 1992; Mini, Palomba, Angrilli, & Bravi, 1996; Palomba, Angrilli, & Mini, 1997; Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Keil et al., 2002; Hajcak, MacNamara, & Olvet, 2010). Such findings indicate that emotionally-laden stimuli attract attention, which might result from the fact that they are automatically perceived and processed as task-relevant.

In addition to electrophysiological indicators of emotional response, as reflected in ERP patterns, other measurements used when examining emotional reactions are based on physiological responses from the autonomic nervous system (ANS). The autonomic nervous system plays a key role in peripheral functions, and involves sympathetic and parasympathetic systems. Importantly, the activity of the ANS encompasses a wide range of functions, including not only emotional processing, but also digestion, homeostasis, effort, attention (Frith & Allen, 1983; Berntson & Cacioppo, 2000; Mauss & Robinson, 2009). The most prominent indicators of the ANS activity are electrodermal and cardiovascular reactions (Mauss & Robinson, 2009). While electrodermal measures include

sweat gland, as reflected in skin conductance level or short-duration skin conductance responses, cardiovascular responses involve heart rate, heart-rate variability, cardiac output, blood pressure, pre-ejection period, and total peripheral resistance, each of such measures reflecting either sympathetic activity, parasympathetic activity, or both (Mauss & Robinson, 2009).

The measurement of the skin conductance level (SCL), also referred to as the galvanic skin response (GSR), is one of the most commonly employed methods of examining ANS activity related to emotional responding to an external stimulus. The galvanic skin response predominantly indexes sympathetic response, and is modulated only by the level of arousal of emotional stimuli. The recorded GSR signal pertains to the resistance between two electrodes that are attached to participants' fingers and connected to a voltage amplifier. The number of galvanic skin responses as well as the magnitude of the GSR amplitude reflects the emotional state of a participant, and is thus employed as a marker of emotional arousal, with a more pronounced GSR response indexing an increased intensity of emotions experienced (Bradley, Cuthbert, & Lang, 1990; Cook, Hawk, Davis, & Stevenson, 1991; Waugh, Thompson, & Gotlib, 2011; Monfort et al., 2014).

Importantly, it has often been assumed that basic emotions vary on the arousal scale. Fear, anger and happiness are usually referred to as high-arousing emotions, while sadness is described as a low-arousing emotion (e.g. Kensinger & Corkin, 2004; Kensinger, 2004). This notwithstanding, previous research has shown that this discrepancy might not necessarily be reflected in galvanic skin response patterns (e.g. Levenson et al., 1990; Kreibig et al., 2007), or even that sadness evokes physiological arousal similar to emotions that are usually referred to as high-arousing (Frazier et al., 2004). This might suggest that there are no distinct relationships between arousal and emotional states experienced at a given point in time (Larsen et al., 2008).

Nonetheless, previous studies into emotional responding have rarely combined the aforementioned research methods, and thus experiments conducted so far have mostly focused on examining either emotional valence or arousal. In our original study (Jankowiak & Korpal, 2018), we examined the processing of emotionally-laden narratives in the native and non-native tongue by means of employing both a self-report measure and the galvanic skin response as an index of the autonomic nervous system. In order to obtain self-report data, we used the SUPIN S30 questionnaire (Brzozowski, 2010), which is a Polish adaptation of the PANAS questionnaire (Watson, Clark, & Tellegen, 1988), and which allowed for examining how participants labeled their emotions. With a view to obtaining valid self-report data, we used an immediate self-response, as the SUPIN questionnaires were administered directly after the presentation of each stimulus. In addition, we employed a GSR measure as a physiological index of emotional arousal to the presented stimuli.

The main objective of the current paper is to discuss main methodological considerations related to the use of physiological and self-report measures in emotion studies, based on our previous research on the processing of emotionally-laden narratives in the native and non-native language, where we employed the SUPIN S30 questionnaire as a self-report tool, and galvanic skin response (GSR) as a physiological measure (Jankowiak & Korpal, 2018). To this aim, we have extended the analyses reported in our original study (Jankowiak & Korpal, 2018) by examining whether there is a correlation between self-report and physiological data, that is whether participants label the intensity of the emotions experienced in line with their physiological responses. The obtained findings are thoroughly discussed in the context of psychological and physiological methods that are employed in emotion studies, pointing to some crucial methodological challenges that are offered by self-report techniques.

Study

In our original experiment (Jankowiak & Korpal, 2018), we examined how bilingual speakers responded to emotionally-laden linguistic stimuli presented in their native (Polish) and non-native language (English) by means of employing the galvanic skin response along with a self-report measure (SUPIN S30). Participants ($N = 27$; 16 females, 11 males) were all late proficient bilingual speakers ($M_{AgeofAcquisition} = 9.07$, $SD = 3.00$), and were students of the Faculty of English at Adam Mickiewicz University in Poznan (Jankowiak & Korpal, 2018, pp. 666–667). Each participant was presented with two narratives in each language (two in the visual modality and two in the auditory modality), which were all aimed to evoke sadness. The experimental materials involved texts concerning death of a relative, and coping with a terminal illness. Participants were asked to read or listen to the stimuli, and to provide answers to yes/no comprehension questions regarding the content of each text (for a more detailed description of materials and procedures applied in the experiment, see Jankowiak & Korpal, 2018, pp. 666–671).

To record skin conductance, we used an ADInstruments GSR Amp galvanic skin response amplifier, adopting 0.02 µS as a minimum limit for measured skin conductance. Two 8 mm diameter reusable electrodes with a silver-silver chloride sensor were attached to medial phalanx of both the index and middle finger of the non-dominant hand of each participant. A 0.5% saline skin resistance/conductance electrode paste was used. Skin conductance was recorded by means of the PsychLab Data Acquisition software, and the collected data were further analyzed by the PsychLab Analysis software. To calculate total response count, all the data were sectioned into blocks of 30 seconds each, and then processed to detect skin conductance responses in each moment of interest. Furthermore, to measure how participants self-reported their emotional states, we employed the SUPIN S30 questionnaire (Brzozowski, 2010), which is the Polish adaptation of PANAS (Positive and Negative Affect Schedule; Watson, Clark, & Tellegen, 1988). Participants were asked to fill in the SUPIN S30 questionnaire directly

after the presentation of text. The analysis performed on the self-ratings obtained from the SUPIN questionnaires was based on values for negatively-valent adjectives only (N = 15). The Negative Affect Score was calculated in line with the scoring instructions (Brzozowski, 2010).

Results

As for the SUPIN results, a 2 modality (visual vs. auditory) × language (Polish vs. English) repeated measures *ANOVA* showed no statistically significant findings, $p > .05$ (Jankowiak & Korpal, 2018, p. 671; Fig. 1). As for the GSR results, a 2 modality (visual vs. auditory) × language (Polish vs. English) repeated measures *ANOVA* revealed an interaction between modality and language, $F(1, 26) = 17.10$, $p < .001$, $\eta_p^2 = .397$. In Polish (the native tongue), texts presented in the visual modality ($M = 7.70$, $SE = .91$) elicited a larger number of skin conductance responses relative to the auditory modality ($M = 6.05$, $SE = .71$), $t = -3.29$, $p = .003$, $r = .294$. Importantly, this difference was not observed in English (the non-native language), where the visual ($M = 5.10$, $SE = .66$) and auditory materials ($M = 5.49$, $SE = .68$) evoked a comparable number of skin conductance responses, $p > .05$ (Jankowiak & Korpal, 2018, p. 671–672; Fig. 2).

Figure 1. Mean SUPIN results for the visual (dark gray) and auditory (light gray) modality in Polish (left-hand side) and English (right-hand side; after Jankowiak & Korpal, 2018: 672)
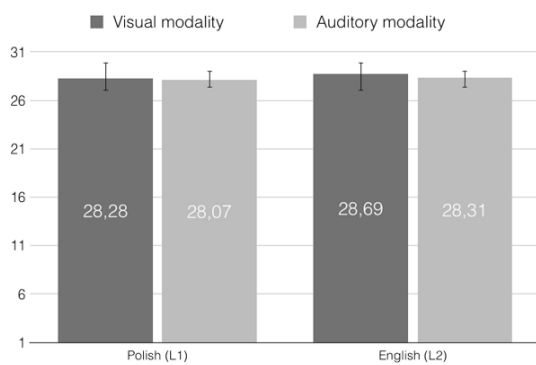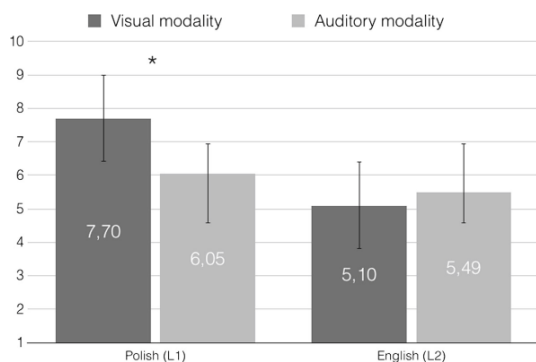

Figure 2. Mean number of galvanic skin responses for the visual (dark gray) and auditory (light gray) modality in Polish (left-hand side) and English (right-hand side; after Jankowiak & Korpal, 2018: 673)

Following the results reported in Jankowiak and Korpal (2018), we decided to further triangulate the two research methods (i.e., SUPIN and GSR) with a view to obtaining more valid results. A correlation analysis was therefore conducted to verify whether GSR values were reflected in how participants self-reported their emotional states. Interestingly, a correlation between a physiological (GSR) and a self-report measure of emotion (SUPIN S30 for negative emotions) was not observed ($rs = .032$; $p > .05$). Hence, an increased physiological arousal in response to emotionally-laden stimuli presented in the native tongue relative to the non-native language was observed only in GSR patterns, and was not reflected in how participants labeled the intensity of the experienced emotions in self-reports. Fig. 3 presents the scatterplot showing the lack of correlation between the number of galvanic skin responses exceeding 0.02 µS and SUPIN scores for negative emotions:

Figure 3. Scatterplot showing the lack of correlation between GSR values and SUPIN scores for negative emotions
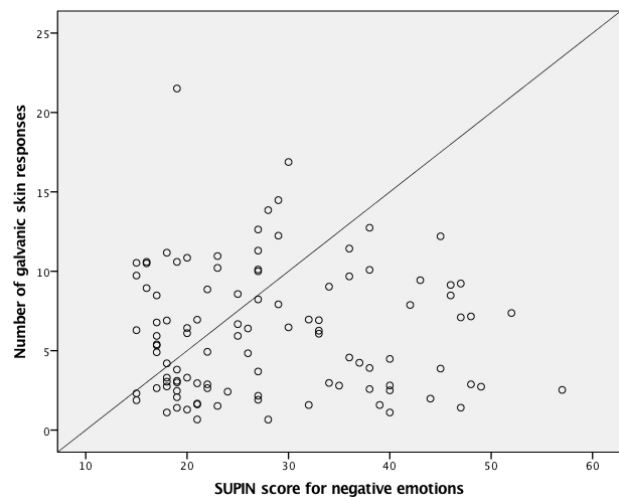
Discussion

The aim of the present paper was to discuss the validity of physiological compared to self-report measures, both of which are frequently employed in order to study human emotional responding. As an extension of the analyses reported by Jankowiak and Korpal (2018), we performed correlational tests between the number of galvanic-skin responses and negative-affect score obtained from self-report data.

Importantly, the results obtained from the study did not show any correlation between the physiological (the number of galvanic skin responses) and a self-report measure of emotion (the SUPIN S30 questionnaire scores). A more pronounced emotional reaction was observed in response to the native relative to the non-native language, which was however reflected only in GSR patterns. No statistically significant between-language differences were, however, observed in SUPIN S30 scores (Jankowiak

& Korpal, 2018). Such results point to the importance of research method triangulation, at the same time raising the question of whether tools based on self-report can be used as valid measures of emotion in an experimental setting. In the following sections, a number of methodological problems of self-report measures will be discussed, followed by more general considerations about the validity of (triangulated) research methods employed to investigate emotional responding.

One of the limitations of self-report research tools, and a plausible explanation for the lack of correlation between emotion markers used in our study, is social desirability bias (Crowne & Marlowe, 1964; Fisher, 1993; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), which refers to "the need for social approval and acceptance and the belief that it can be attained by means of culturally acceptable and appropriate behaviors" (Crowne & Marlowe, 1964, p. 109). Social desirability bias is thus triggered by "the desire of respondents to avoid embarrassment and project a favorable image to others" (Fisher, 1993, p. 303). In emotion studies, participants might not be willing to give true answers to questions about their current emotional states, as they want to present a positive image in front of the experimenters (Welte & Russell, 1993). For example, participants may either conceal their emotions or pretend that they experience emotions, which in fact they fail to experience. In our study, participants were presented with four experimental emotionally-laden narratives, two of which were presented in their non-native tongue, which – due to a decreased sensitivity to emotional materials in the non-native language – should have been perceived as evoking a decreased emotional response relative to the native stimuli. Yet, as participants were aware that their emotional responding was tested in the experiment, their reactions might have been affected by social desirability bias, as a result of which they self-reported their current emotional state to be affected emotionally by all four emotionally laden stimuli, irrespectively of language nativeness. As a consequence, no statistically significant difference in self-reported emotional states with reference to the native and non-native narratives was observed.

Another limitation of using self-report measures in emotion studies pertains to the fact that some items included in questionnaires may be regarded by participants as sensitive questions, which are considered intrusive and are assumed to demonstrate a threat of disclosure (Tourangeau & Yan, 2007). Similar to social desirability bias, this would result in providing invalid self-reports of participants' current emotional states. For example, the original English version of PANAS (Positive and Negative Affect Schedule) contains items such as: *guilty*, *hostile*, *scared*. This also applies to SUPIN S30, the Polish adaptation of PANAS by Brzozowski (2010), which was employed in our study. Although the items do not refer to participants' personal, or intimate experiences, using the questionnaire may at least to some extent lead to personal self-disclosure. As a consequence, participants may be unwilling to give true answers to the questions.

The widely used argument in favor of employing self-reported measures to study emotions pertains to the fact that such measures enable participants to recognize and label their emotions (Russell & Barrett, 1999). However, in fact some participants may be unaware of the fact that they experience a given emotion, especially in the ecologically non-valid experimental setting (cf. alexithymia, Lane, Ahern, Schwartz, & Kaszniak, 1997). In such an event, the physiological arousal will not be reflected in the participant's self-assessment of their emotional states. Apart from being unable to recognize their own emotions, participants can also falsely categorize them, as well as fail to identify the sources of emotion, or the relationship between emotions and behavior (Reykowski, 1968). All these distortions may have a crucial influence on mean self-report scores. To mitigate this problem, we used an immediate self-report response, which is considered to be more valid than a delayed response (Robinson & Clore, 2002; Mauss & Robinson, 2009). Nevertheless, it is still possible that some participants failed to recognize their sadness, which might have resulted in the lack of correlation between the activation of the sympathetic nervous system and self-reported emotional states.

The aforementioned methodological problems all refer to the limitations of self-reported measures. Nevertheless, it should also be noted that methods based on measuring physiological arousal have also been criticized, as they are believed to compromise the study's ecological validity. Although it is not as obtrusive as some other methods used to investigate arousal and stress (e.g., concentration of blood cortisol), using GSR can be problematic due to electrodes placed on the participant's fingers, which may restrict their movements during the experiment. In this way, the study's ecological validity is compromised, as the experiment fails to approximate the real world and, thus, participants' reactions might differ from their behavior in their natural environment. Nevertheless, when compared to self-report psychological instruments, physiological arousal appears to be a more valid and less biased method of studying momentary emotional states in an experimental setting.

One of the most significant challenges in experimental research on emotions is that some physiological arousal might be generated in participants, as they are aware that they take part in an experiment and their performance is evaluated. Sesen and Mutlu (2014) came up with the idea of laboratory anxiety, which assumes that some participants might be afraid to take part in experiments. Laboratory anxiety may be triggered by the sense of uncertainty about the experimental procedure. Similarly, participants may be physiologically aroused as a result of test anxiety, or fear of negative evaluation (Horwitz, Horwitz, & Cope, 1986; Du, 2009). Although in the study by Jankowiak & Korpal (2018) participants did not perform any task to which formative assessment would be provided, it is possible that their answers were affected by the fear of being negatively assessed by the experimenter.

Conclusion

The present article was aimed to discuss various methods to study emotional responding, with a special focus on self-report and physiological measures, often employed in emotion studies. Using the example of the study conducted on Polish-English bilingual speakers (Jankowiak & Korpal, 2018), in which we observed discrepant results for GSR values and SUPIN S30 scores, we have attempted to discuss potential methodological considerations of measuring individuals' emotional reactions in an experimental setting. The idea behind research method triangulation is that using more than one measurement tool allows for obtaining more valid data. However, the analysis of the results of our study might point to potential limitations of using a self-reported measure in experimental research aimed at investigating emotional responding, including social desirability bias, sensitive questions, and lack of emotional self-awareness. On the other hand, self-report measurements are often employed as, unlike methods employed in order to examine the activation of the autonomic nervous system, they provide information not only on emotional arousal, but also on the valence of the presented stimuli. The present article offers only a tentative insight into the question of the validity of self-report measures in emotion studies. More research is needed to have a better understanding of their limitations, identify the sources of method biases, and to provide recommendations for how to mitigate the problems involved in studying emotional responding.

Acknowledgements

References

Berntson, G. G., & Cacioppo, J. T. (2000). From homeostasis to allodynamic regulation. In: J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 459–481). New York: Cambridge University Press. doi: https://doi.org/10.1017/9781107415782.018

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1990). Startle reflex modification: Emotion or attention? *Psychophysiology*, *27*(5), 513–522.

Brown, S. B. R. E., van Steenbergen, H., Band, G. P. H., de Rover, M., & Nieuwenhuis, S. (2012). Functional significance of the emotion-related late positive potential. *Frontiers in Human Neuroscience*, *6*(33). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3287021/. doi: http://dx.doi.org/10.3389/fnhum.2012.00033

Brzozowski, P. (2010). *Skala Uczuć Pozytywnych i Negatywnych (SUPIN). Polska adaptacja skali PANAS Davida Watsona i Lee Anny Clark. Podręcznik* [*SUPIN – Polish adaptation of PANAS – Positive and Negative Affect Schedule by David Watson and Lee Anna Clark. A manual*] Warsaw: Pracownia Testów Psychologicznych.

Cacioppo, J., Berntson, G. G., Klein, D. J., & Poehlmann, K. M. (1997). The psychophysiology of emotion across the lifespan. *Annual Review of Gerontology and Geriatrics*, *17*, 27–74.

Cohen, M. X. (2014). *Analyzing neural time series data.* Cambridge: MIT Press.

Cook, E. W., Hawk, L. H., Davis, T. L., & Stevenson, V. E. (1991). Affective individual differences and startle reflex modulation. *Journal of Abnormal Psychology*, *100*, 5–13. doi: http://dx.doi.org/10.1037/0021-843X.100.1.5

Crowne, D., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*, 95–111. doi: https://doi.org/10.1016/S0301-0511(99)00044-7

Davidson, R. J. (1999). Neuropsychological perspectives on affective styles and their cognitive consequences. In: T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 103–123). New York: Wiley. doi: 10.1002/0470013494.ch6

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. doi: http://dx.doi.org/10.1111/1467-9280.00221

Du, X. (2009). The affective filter in second language teaching. *Asian Social Science 5*(8), 162–165. doi: http://dx.doi.org/10.5539/ass.v5n8p162

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*, 456–467. doi: 10.1111/j.1469-8986.1977.tb01312.x

Ekman, P. (1999). Basic emotions. In: T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). New York: Wiley. doi: http://dx.doi.org/10.1002/0470013494.ch3

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research 20*(2), 303–315. doi: http://dx.doi.org/10.1086/209351

Frazier, T. W., Strauss, M. E., & Steinhauer, S. (2004). Respiratory sinus arrhythmia as an index of emotional response. *Psychophysiology 41*, 75–83. doi: https://doi.org/10.1046/j.1469-8986.2003.00131.x

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358. doi: http://dx.doi.org/10.1037/0003-066X.43.5.349

Frith, C. D., & Allen, H. A. (1983). The skin conductance orienting response as an index of attention. *Biological Psychology*, *17*, 27–39. doi: http://dx.doi.org/10.1016/0301-0511(83)90064-9

Gottman, J. M., & Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. *Journal of Consulting and Clinical Psychology*, *53*(1), 151–160. doi: http://dx.doi.org/10.1037/0022-006X.53.2.151

Gross, J. J. (2007). *Handbook of emotion regulation*. New York: Guilford Press.

Hajcak, G., & Olvet, D. M. (2008). The persistence of attention to emotion: Brain potentials during and after picture presentation. *Emotion*, *8*, 250–255. doi: 10.1037/1528-3542.8.2.250

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*(2), 129–155. doi: http://dx.doi.org/10.1080/87565640903526504

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. In: E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 27–36). Englewood Cliffs, NJ: Prentice Hall.

Jankowiak, K., & Korpal, P. (2018). On modality effects in bilingual emotional language processing: Evidence from galvanic skin response. *Journal of Psycholinguistic Research*, *47*, 663–677. doi: https://doi.org/10.1007/s10936-017-9552-5

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, *39*, 641–649. doi: http://dx.doi.org/10.1111/1469-8986.3950641

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neuroscience*, *15*, 241–251. doi: https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. doi: https://doi.org/10.1073/pnas.0306408101

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear-

and sadness-inducing films. *Psychophysiology*, *44*, 787–806. doi: 10.1111/j.1469-8986.2007.00550.x

Kutas, M., Van Petten, C. K., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In: M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics: 2nd Edition* (pp. 659–724). San Diego: Academic Press.

Lane, R. D., Ahern, G. L., Schwartz, G. E., & Kaszniak, A. W. (1997). Is alexithymia the emotional equivalent of blindsight? *Biological Psychiatry*, *42*(9), 834–844. doi: http://dx.doi.org/10.1016S0006-3223(97)00050-4

Lang, P. J. (1988). What are the data of emotion? In: V. Hamilton, G. H. Bower & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 173–191). New York: Kluwer Academic/Plenum Publishers. doi: http://dx.doi.org/10.1007/978-94-009-2792-6

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In: P. J. Lang, R. F. Simons & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97–135). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Larsen, R. J., & Prizmic-Larsen, Z. (2006). Measuring emotions: Implications of a multimethod perspective. In: M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 337–351). Washington, DC: American Psychological Association. doi: http://dx.doi.org/10.1037/11383-000

Larsen, R. J., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In: M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of emotions (third edition)* (pp. 180–195). New York: Guilford Publications, Inc.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford, UK: Oxford University Press.

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, *60*, 587–597. doi: http://dx.doi.org/10.1037/0022-3514.45.3.587

Levenson, R. W., Ekman, P., & Friesen, W. F. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, *27*(4), 363–384. doi: http://dx.doi.org/10.1111/j.1469-8986.1990.tb02330.x

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*(2), 209–237. doi: http://dx.doi.org/10.1080/02699930802204677

Mini, A., Palomba, D., Angrilli, A., & Bravi, S. (1996). Emotional information processing and visual evoked brain potentials. *Perceptual and Motor Skills*, *83*, 143–152. doi: http://dx.doi.org/10.2466/pms.1996.83.1.143

Monfort, S. S., Kaczmarek, Ł. D., Kashdan, T. B., Drążkowski, D., Kosakowski, M., Guzik, P., Krauze, T., & Gracanin, A. (2014). Capitalizing on the success of romantic partners: A laboratory investigation on subjective, facial, and physiological emotional processing. *Personality and Individual Differences*, *68*, 149–153. doi: https://doi.org/10.1016/j.paid.2014.04.028

Naumann, E., Bartussek, D., Diedrich, O., & Laufer, M. (1992). Assessing cognitive and affective information processing functions of the brain by means of the late positive complex of the event-related potential. *Journal of Psychophysiology*, *6*, 285–298.

Palomba, D., Angrilli, A., & Mini, A. (1997). Visual evoked potentials, heart rate responses and memory to emotional pictorial stimuli. *International Journal of Psychophysiology*, *27*, 55–67. doi: http://dx.doi.org/10.1016/S0167-8760(97)00751-4

Panksepp, J. (2007). Neurologizing the psychology of affects: How appraisal-based constructivism and basic emotion theory can coexist. *Perspectives on Psychological Science*, *2*(3), 281–295.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology 88*(5), 879–903. doi: 10.1037/0021-9010.88.5.879

Reykowski, J. (1968). *Eksperymentalna psychologia emocji* [*Experimental psychology of emotion*]. Warsaw: Książka i Wiedza.

Robinson, M. D., & Clore, G. L. (2002). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, *83*(1), 198–215. doi: 10.1037//0022-3514.83.1.198

Ruef, A. M., & Levenson, R. W. (2007). Continuous measurement of emotion. In: J. A. Coan, & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 286–297). Oxford: Oxford University Press.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. doi: http://dx.doi.org/10.1037/0022-3514.76.5.805

Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality & Social Psychology*, *5*, 37–63.

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. In: S. Anders, G. Ende, M. Junghöfer, J. Kissler, & D. Wildgruber (Vol. Eds.), *Progress in brain research: Vol. 156. Understanding emotions* (pp. 31–51). Amsterdam: Elsevier. doi: http://dx.doi.org/10.1016/S0079-6123(06)56002-9

Sesen, B. A., & Mutlu, A. (2014). An action research to overcome undergraduates' laboratory anxiety. *Procedia – Social and Behavioral Sciences*, *152*, 546–550.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*(4), 813–838. doi: http://dx.doi.org/10.1037/0022-3514.48.4.813

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883. doi: http://dx.doi.org/10.1037/0033-2909.133.5.859

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838.

Waugh, C. E., Thompson, R. J., & Gotlib, I. H. (2011). Flexible emotional responsiveness in trait resilience. *Emotion*, *11*(5), 1059–1067. doi: http://dx.doi.org/10.1037/a0021786

Welte, J. W., & Russell, M. (1993). Influence of socially desirable responding in a study of stress and substance abuse. *Alcoholism: Clinical and Experimental Research*, *17*(4), 758–761. doi: 10.1111/j.1530-0277.1993.tb00836.x