Transportation Research Part F: Psychology and Behaviour 83 (2021) 130–147

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

The role of feedback in shaping responses to risky road scenarios: Evidence from electrodermal activity

Evelyn Gianfranchi [a,*], Mariaelena Tagliabue [b,c], Alberto Megías-Robles [a,d], Antonio Maldonado [a]

[a] *Mind, Brain, and Behavior Research Center-CIMCYC, University of Granada, Campus de Cartuja s/n, Granada, Spain*
[b] *Department of General Psychology, University of Padova, Via Venezia 8, Padova, Italy*
[c] *Department of Civil, Environmental and Architectural Engineering, University of Padova, Via Marzolo 9, Padova, Italy*
[d] *Department of Basic Psychology, University of Malaga, Campus de Teatinos s/n, Malaga, Spain*

ARTICLE INFO

Keywords:
Feedback
Risky driving
Electrodermal activity
Risk perception
Decision-making

ABSTRACT

When faced with hazardous driving situations, rapid and effective risk perception and decision-making processes are of crucial importance for avoiding crashes. In these cases, the processes are accompanied and influenced by underlying psychophysiological mechanisms such as electrodermal activity (EDA) modulations. The present work aims to assess the psychophysiological mechanisms underlying participants' risk perception and decision-making when facing risky road scenarios, as correlated to the feedback role in modulating participants' behavior. Study 1 ($n = 32$) explores the behavioral effects of administering a contingent feedback in a decision (decision-making) and an evaluation (risk perception) task in response to a set of risky and not risky images. The behavioral data reveal an effect on the participants' probability of response, independently from the type of image presented, when the feedback was administered. In the decision task, the effect is accompanied by a change in the amplitude and percentage of the skin conductance responses (SCRs), which are moderated by block of trials. Study 2 ($n = 44$) better assesses the role of task and block on participants' physiological activation, as measured by EDA signal. The results show an increase in psychophysiological activation when the feedback is delivered, in the first part of the tasks, both in terms of SCRs amplitude and percentage to the presented road scenarios, followed by a decrease in the second part of the tasks. Moreover, this effect is more evident in the decision task than the evaluation task. These findings suggest that the role exerted by feedback when facing risky traffic images may be described as based on an associative process that, once the correct response has been learned, tends to be reduced as it becomes automatic. Overall, the results of the two studies represent an important step toward the development of training programs aimed at promoting safer behaviors in risky driving contexts.

1. Introduction

In daily life, everyone must face situations that require rapid and effective risk perception (*i.e.,* the subjective judgment of the nature and severity of a potential risk) and decision-making. This is particularly true in complex and risky contexts, such as when

* Corresponding author at: Mind, Brain, and Behavior Research Center – CIMCYC, University of Granada, Campus de Cartuja s/n, Granada, Spain.
E-mail address: egia@ugr.es (E. Gianfranchi).

https://doi.org/10.1016/j.trf.2021.10.001
Received 6 February 2021; Received in revised form 29 July 2021; Accepted 5 October 2021
Available online 25 October 2021
1369-8478/

driving on the road. In this case, ineffective risk perception and erroneous decision-making can easily lead to crashes and the attendant consequences (Megías, Maldonado, Cándido, & Catena, 2011; Ulleberg & Rundmo, 2003). Given the importance of these processes, various models have attempted to explain how people perceive risk and make decisions in complex contexts (*e.g.,* Bechara, Damasio, Damasio, & Lee, 1999; Damasio, 1994; Bechara, Tranel, Damasio, & Damasio, 1996; Epstein, 1994; Kahneman & Frederick, 2005; Sanfey & Chang, 2008; Slovic, Finucane, Peters, & MacGregor, 2007). Most of these have theorized the presence of a ***dual process*** underlying risk perception and decision-making, that is, the involvement of two separate systems at different levels, depending on the characteristics of the situation at hand. Thanks to the interplay between these two systems, we would be able to process the surrounding environment and to make decisions that are appropriate to each specific situation (Megías et al., 2015; Reyna, 2004; Slovic, Finucane, Peters, & MacGregor, 2004).

1.1. The basis of the dual process models

Among the aforementioned ***dual process*** theories (see Osman, 2004 for a review), that aim at explaining how people perceive risk and how they react to it, through decision-making, two models represent the core of the literature on this specific topic, namely the dual process model (Epstein, 1994), and the affect heuristic theory (Slovic et al., 2007). Both models postulate the existence of two distinct systems (*i.e.*, cognitive and physiological mechanisms), that differently and jointly regulate human decision-making when facing risky situations.

Although each theory describes the two systems with exact terminologies and define features that are specific to each of them, they generally agree on the distinction between a system that is more automatic, associative, rapid, action-oriented (from now on, experiential-affective) (Epstein, 1994) and one that is more controlled, conscious, slow, processing-oriented (from now on, rational-analytic) (Epstein, 1994; Kahneman, 2011; Megías et al., 2015; Slovic et al., 2007).

While the first system produces rapid and effective responses with a low cognitive effort, the second system requires much more cognitive effort and time to process information, and therefore, it is not always applicable in risky and complex situations, when a more rapid response is required, such as during driving. In this type of situations, the experiential-affective system would guide the decision-making toward the most rapid, and possibly efficient, response, as predicted by the main dual process models (Bechara et al., 1996, 1999; Damasio, 1994; Megías et al., 2011; Slovic et al., 2007).

1.2. Psychophysiological modulations correlated to the experiential-affective system

It should be noted that the experiential-affective system appears to be accompanied by underlying psychophysiological modifications (*e.g.,* the so-called somatic marker; Bechara et al., 1996, 1999; Damasio, 1994, Kinnear, Kelly, Stradling, & Thomson, 2013; Slovic et al., 2007), expressed, among others, by changes in electrodermal activity (EDA), that is "changes in the electrical activity of palmar and plantar skin, being concomitants of psychological phenomena" (Boucsein et al., 2012; p. 1017).

To illustrate, the mental representation of a stimulus or a situation would be associated with a "label" of its emotional component and of the related physiological modifications, as coded through previously experienced outcomes. If this mental representation has been linked to a negative emotional and physiological component, when facing the same or a similar situation, a general "state of alarm" would be generated, involving specific psychophysiological modulations (somatic marker), such as changes in certain features of the EDA (*e.g.*, in the so-called skin conductance responses [SCR; *i.e.*, increase of at least 0.05 μS in the EDA signal]). This general state of alarm would then guide decision-making in response to that situation (*e.g.,* by preventing a potentially dangerous behavior) (Bechara et al., 1996, 1999; Bechara & Damasio, 2005).

1.2.1. Somatic marker and driving behavior

To date, a number of studies have investigated the role of the somatic marker in risk perception and decision-making in driving contexts as a function of driving experience or training. For instance, experienced drivers have been reported to show greater percentages of SCRs than novice drivers when exposed to video clips displaying risky road scenes, thus providing evidence of greater activation of the somatic marker (Kinnear et al., 2013).

A series of studies has also demonstrated that virtual training for novice drivers on a moped simulator—where participants could experience virtual crashes as a consequence of their unsafe driving behaviors—was able to elicit a greater emotional involvement (in terms of SCRs) than a passive vision of video clips reproducing risky road scenes (Tagliabue & Sarlo, 2015). Moreover, when the participants drove the moped simulator along the same virtual scenarios twice in two different sessions, earlier SCRs were elicited during the second session than during the first (Gianfranchi, Sarlo, & Tagliabue, 2017; Tagliabue, Gianfranchi, & Sarlo, 2017; Tagliabue, Sarlo, & Gianfranchi, 2019). Finally, an increased percentage of SCRs was found for the hazardous situations in which the participants had a crash or a near miss, in comparison with the situations in which the participants' driving behavior was safe (Tagliabue et al., 2019), indicating a greater activation of the somatic marker.

On the other hand, another study (Megías, Cortés, Maldonado, & Cándido, 2017) showed an increase in safe driving performance on a moped simulator among participants who first received a specific PC training. Throughout the training task (*i.e.,* deciding whether to brake or not given a set of images representing various traffic situations, either risky or not risky), negative feedback (*i.e.,* pictures of real accidents) was delivered to the participants in the experimental group on 50% of the trials in which they decided not to brake when presented with a risky image. However, since no psychophysiological metrics were included in this study, it is not possible to draw any conclusions regarding the presence and modulation of psychophysiological mechanisms underlying the participants' behavior and any possible feedback effect.

1.3. The role of feedback and its unclear underpinning mechanisms

The role of feedback in shaping driving behaviors has been widely studied in the last decade (Aidman, Chadunow, Johnson, & Reece, 2015; Farah et al., 2014; Kluger & De Nisi, 1996; Maldonado, Torres, Catena, Cándido, & Megías-Robles, 2020).

For instance, the use of contingent negative feedback (that is, feedback contingently applied to the behavior) has consistently been shown to promote correct responses in a computerized behavioral task in which the participants had to decide whether or not to brake when faced with a set of traffic images (Maldonado, Serra, Catena, Cándido, & Megías, 2016; Maldonado et al., 2020; Torres, Megías, Catena, Cándido, & Maldonado, 2017). This task, also called "Decision task", is thought to bring into play the activation of the experiential-affective system, due to the urgency of the required response that would make less possible the (slower) activation of the other system (rational-analytic). In this task, the contingent negative feedback appears to modify the individual decision criterion (that is, the response bias) toward a better perception of risk, leading to risk-averse behaviors (Maldonado et al., 2016, Torres et al., 2017). Conversely, the use of non-contingent feedback was found to generate the opposite effect, that is, a reduction in the frequency of braking decisions (Torres et al., 2017).

However, the effects of a contingent negative feedback are not so immediate when delivered in a different type of task, the so-called "Evaluative task". In this task, participants have to rate the images as risky or not risky (thus increasing the influence of the rational-analytic system). When this task has been used to study the effect of feedback in comparison to the Decision task, although the effect is still evident, the participants were generally more rapid in responding during the Decision task when compared with the performance at the Evaluative task (Maldonado et al., 2016). This suggests that the types of required responses in the two tasks are differentially controlled by the two systems proposed by the dual-process theories (Maldonado et al., 2016). In the same vein, the behavioral effect of feedback on both the urgent and evaluative tasks seems to be paralleled by system-specific alterations in brain activation (Megías et al., 2015), which provides further evidence of the involvement of two different systems (Megías et al., 2015).

Overall, the behavioral effect of feedback on both tasks (and therefore, both systems) could be explained in two ways. For instance, feedback could exert its effect on attentional processes involved (Donmez, Boyle, & Lee, 2007; Torres et al., 2017). In this case, contingent feedback could enhance attention towards the salient features of the stimuli, that is the risky images, thus reducing the tendency for participants to be distracted throughout the task, which would improve and speed up their responses. Another possibility is that feedback could have an effect on associative learning and memory processes (Feng & Donmez, 2013; Torres et al., 2017). In this case, the behavioral feedback effect would be a consequence of a series of associations between the stimulus, incorrect behavior, and the corresponding penalty. Thus, contingent negative feedback would induce safer behaviors. This latter explanation is compatible with the somatic marker hypothesis, which postulates an association between the features of a stimulus, the outcomes of previous behaviors and the underlying psychophysiological modulations (such as EDA) that drive the current decision-making process (Bechara et al., 1996, 1999; Bechara & Damasio, 2005).

However, to the best of our knowledge, no studies have yet explored the existence of a possible link between the effect of feedback and the dual-process models, such as the somatic marker hypothesis, in traffic scenarios, and there have been no attempts to assess the psychophysiological modulations underlying the feedback effect in terms of EDA. Thus, the aim of the present work was to address these two issues.

1.4. The present study

As a first step, we designed a three-stage work with a threefold aim.

The **first aim** was to replicate the previous behavioral results observed in the literature regarding the effect of contingent negative feedback in both urgent tasks (from now on, Decision; the participant must decide whether or not to brake when facing pictures of various road scenarios) and evaluative (from now on, Evaluation; the participant must evaluate whether or not a road scenario is risky) tasks in an Italian sample, exploring potential cultural differences in these effects. To this end, we designed an experiment (Study 1-phase A) to replicate, in an Italian sample, previous findings regarding an effect of feedback on participant's probability of response, that is, an increased probability of response when facing traffic scenarios, either risky or not (*e.g.*, Torres et al., 2017). Indeed, previous studies were always performed on young Spanish populations (university students; *e.g.*, Maldonado et al., 2020; Torres et al., 2017), leading to the evidence that the use of a contingent negative feedback in this population produces consistent changes in participants' decision-making, toward more risk-averse behaviors (*i.e.*, when facing traffic images, deciding more frequently to brake/to evaluate them as risky). Moreover, as previously noted, these changes seem to transfer to the participants' driving behavior on a simulator, leading to safer behaviors (Megías et al., 2017). To date the only study that assessed the possible cultural differences linked to similar tasks with the same set of images employed in the present study (four samples from Ukraine, Italy, Spain, and Sweden respectively; Di Stasi et al., 2020), found that cultural factors seem to modulate the general participants' promptness to respond and visually explore the images. However, since the study did not use any feedback, it is not possible to draw any conclusion regarding the cognitive and psychophysiological mechanisms related to the effect of the feedback and, if any, its cultural modulations.

Therefore, the **second aim** of the work was to assess whether (and how) it is possible to identify stimulus-related EDA modulations in terms of amplitude and frequency of SCRs, possibly correlated to contingent negative feedback. To this end, we designed an experiment that focused on the identification of the psychophysiological correlates of the contingent negative feedback (the EDA signal) during a Decision task (Study 1 - phase B). If present, these moderations should appear in the EDA signal in response to the stimuli (the road scenarios), leading to an increased amplitude and/or percentage of SCRs elicited by the stimulus in the conditions when the feedback is administered than when it is not. During phase B, we decided to focus on the Decision task given the evidence that the EDA modulations are frequently shown in situations where the participants are forced to quickly react to traffic hazards (*e.g.*, Kinnear et al., 2013; Tagliabue, Gianfranchi, & Sarlo, 2017, 2019; Tagliabue & Sarlo, 2015).

However, since the comparison between the Decision and the Evaluation task is key, in the third step we designed a new study to compare the EDA signal moderations in the two tasks (Study 2). It was hypothesized that the presence of a contingent negative feedback should be reflected in participants' behavior, that is, an increased probability of response when facing risky traffic images compared with participants that did not receive any feedback, particularly in the case of the Decision task. Moreover, these effects should be paralleled by an increase in physiological activation in terms of increased amplitude and/or percentage of SCRs elicited by the stimulus for the participants that received feedback. Finally, this last experiment was run again on a Spanish sample, allowing to explore once again possible cultural specificities (first aim of the work) in these mechanisms, if present.

Taken together, our findings should help to disentangle the two explanations regarding the way in which contingent feedback is supposed to exert its effect on the participants' behavior, that is, through either an attentional or an associative process (**third aim** of the work). Moreover, the possible identification of differences between the Italian and the Spanish samples, maybe also in terms of psychophysiological activation, would represent the first step toward the possibility of develop *ad-hoc* training methods to improve drivers' risk perception, taking into account cultural differences when needed.

2. Study 1

2.1. Participants

We used a convenience sampling method to recruit thirty-two Italian students from the University of Padova, that agreed to take part in the study (mean age ± standard deviation [SD] = 21.47 ± 2.26 years; age range: 19–28 years; 20 females). Sixteen of the participants were assigned to the Feedback group (mean age ± SD = 21.00 ± 2.19 years; age range: 19–28 years; 11 females; 14 holding a valid car driving license) whilst the remainder were assigned to the Control group (mean age ± SD = 21.93 ± 2.29 years; age range: 19–28 years; 9 females; all holding a valid car driving license). All participants had normal or corrected to normal vision, did not report any neurological disorder or use of medications. They did not receive any compensation for their participation in the study. All the participants gave informed consent and were assured of their rights according to the Code of Ethics of the World Medical Association (WMA) (Declaration of Helsinki – WMA, 2008), under which suggestions the study was conducted.

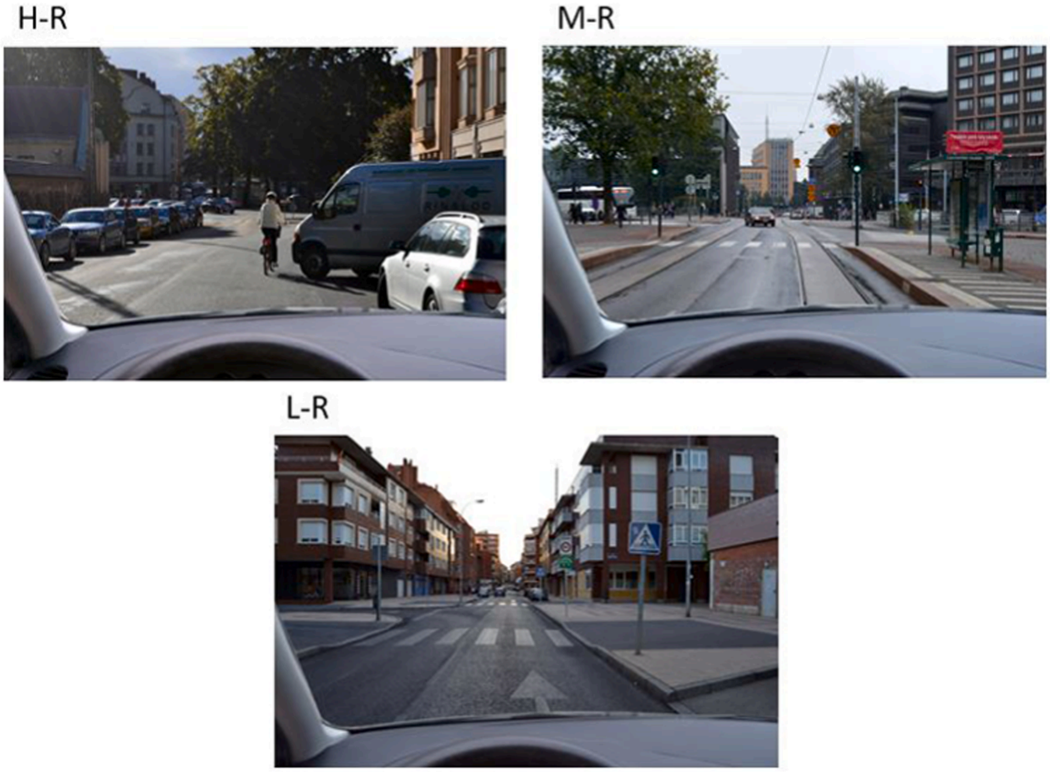

Fig. 1. An example of the traffic scenarios employed in the two studies, namely a high-risk image (H-R, upper left panel), a medium-risk image (M-R, upper right panel), and a low-risk image (L-R, bottom panel).

2.2. Stimuli

One hundred and twenty images showing actual first-person traffic situations were employed in this study (see Fig. 1). The images were selected from an extensive database of traffic scenario pictures taken on Spanish and Finnish roads, which have previously been used to evaluate levels of perceived risk (see Megías, Cándido et al., 2018). Sixty of the selected pictures depicted **high-risk situations** (average risk = 4.56 [0: no risk; 7: high risk]) and the other sixty showed no risk (average risk = 1.65). The risky situations were chosen so that to avoid the risk, the best option was always to brake. For the **low-risk situations**, we ensured that braking did not make sense, excluding those situations such as red traffic lights or give way and stop signs, where braking is a possibility (or necessity) although there is no imminent risk to avoid (that were classified as **medium-risk**). Moreover, all the selected images met a set of statistical criteria in order to reduce the interpersonal variability in the interpretation of the traffic scenario (see Megías, Cándido et al., 2018 for further details). Specifically, all the images included in the original database were evaluated by forty driving school instructors in relation to the perceived speed and the best response needed to avoid the hazard. We considered only those images where the mean standard deviation of the perceived speed was lower than 25% and the best option to avoid the hazard (if any) was to brake (for at least 70% of the driving instructors).

During phase A, each image was displayed twice in each task (see Section 2.4), producing a total of 240 trials in each task. The images were displayed through the Eye-Link Software Experiment Builder (SR Research Ltd., Ottawa, Ontario, CA – version 1.10) on a Windows 7 system PC, with a 14-inch monitor and a resolution set to 1024 × 768, placed at a distance of ≈75 cm from the participant. During phase B, a subset of the 120 images included in phase A was employed, obtained by selecting 20 images from those considered as high-risk (average risk = 4.50) and 20 from those considered to show no risk (average risk = 2.12), for a total of 40 trials in each condition (that is, with and without feedback) (see above for further details regarding stimuli selection). The images were displayed through E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA – version 2.0.10) using the same PC.

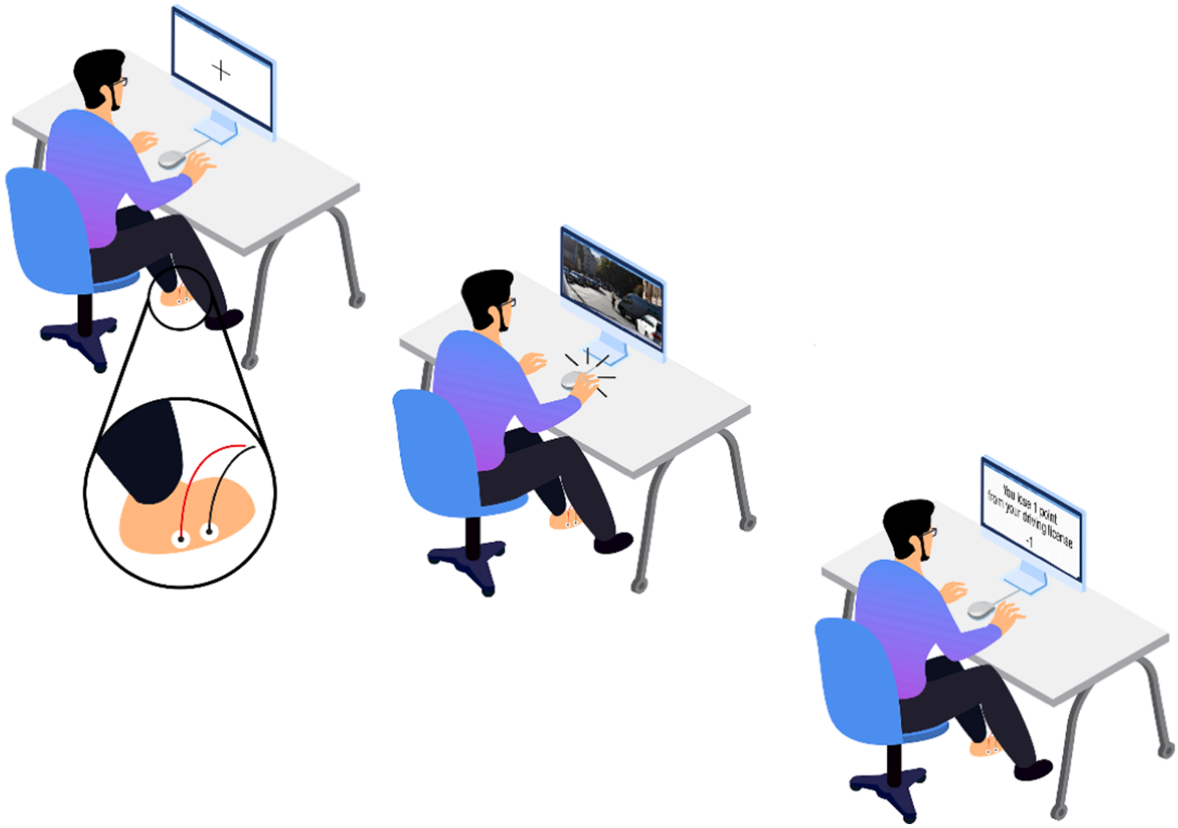

Fig. 2. Schematic representation of the main experimental setting features along the studies. In all the three experiments (Study 1 – phase A and B; Study 2) the participants underwent a series of trials presented on a PC screen while seating in front of it at a distance of about 75 cm (first panel on the top left). Each trial was preceded by a fixation cross of variable duration depending on the experiment. The participants' task was to decide whether or not to brake/evaluate the presented images as risky or not, by pressing a mouse button (central panel). After their answer, depending on the experiment and on the condition, they were presented with a negative feedback, a neutral message, or a blank screen (third panel on the bottom right), again with variable duration depending on the experiment. In Study 1- phase B and in Study 2 the participants' EDA signal was recorded along the entire task using two superficial electrodes placed on the inner side of the left foot (black circle below the first panel).

2.3. EDA recording and processing

During phase B, the participants' EDA was recorded by employing two Ag/AgCl electrodes (1 cm diameter), filled with K-Y lubricating jelly. In order to prevent the recording of any movement artifact due to the need to use both hands to perform the task (see Section 2.4), the electrodes were placed on the left foot over the abductor hallucis muscle—adjacent to the sole of the foot and midway between the proximal phalanx of the big toe, on a point directly beneath the ankle (see Boucsein et al., 2012). A Grass Model SCA1 skin conductance coupler, associated to a Grass CP122 AC/DC Strain Gage amplifier (Grass Instrument Co., W. Warwick, RI, USA) that provided a 0.5-V constant voltage across electrodes, was set to a low-pass 10 Hz filter. The amplifier was connected to a Windows 7

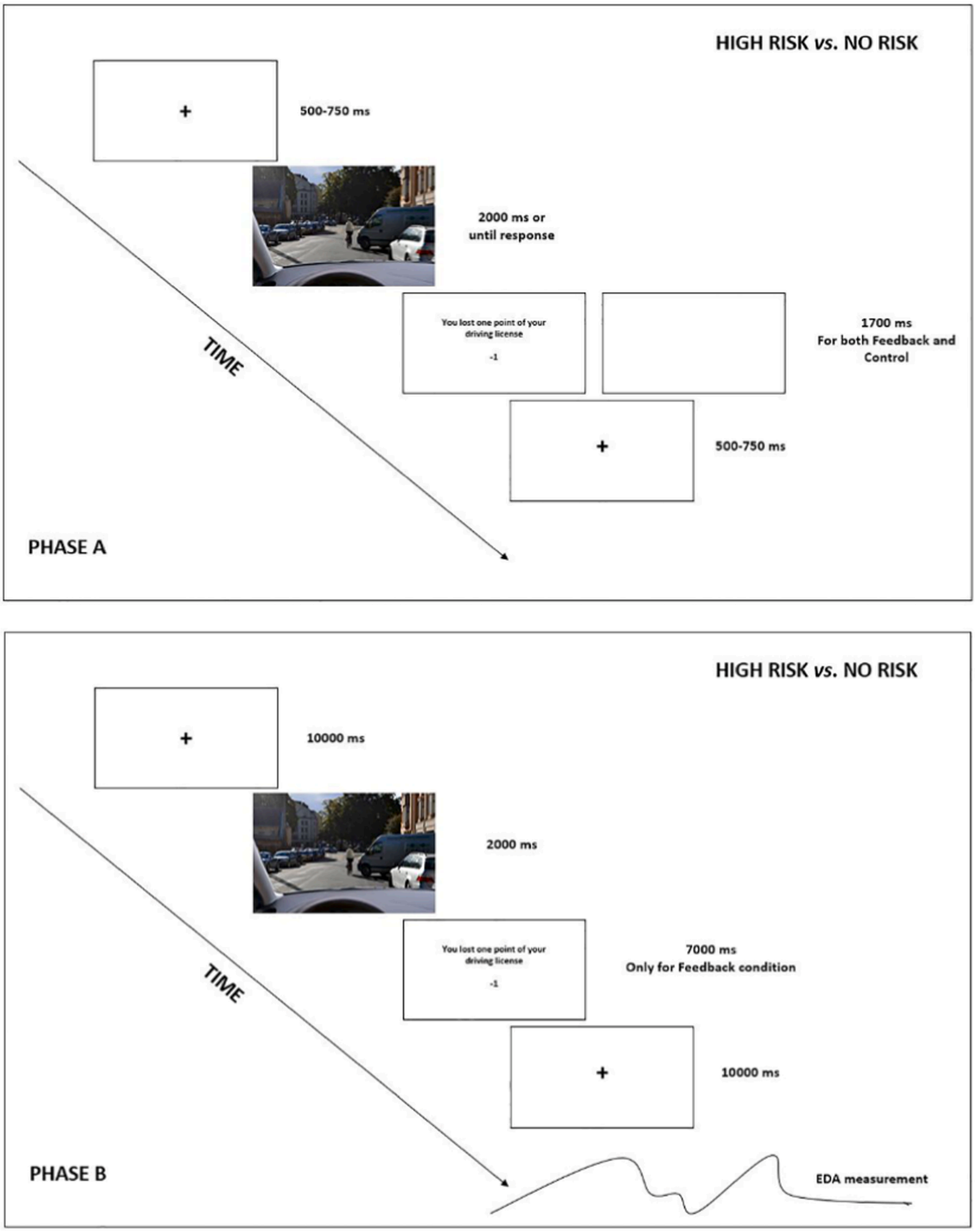

Fig. 3. Schematic representation of a trial for the tasks included in phase A (upper panel) and phase B (lower panel) of Study 1. Throughout both conditions of phase B, we measured the participants' EDA (represented by a curved line resembling the typical shape of the signal at the bottom of the lower panel).

system PC running the Anscovery software (SparkBio Srl, Bologna, Emilia Romagna, IT) for the recording of the EDA. The system recorded and displayed on a dedicated monitor, not visible to the participants, the signal and the markers corresponding to the different stimuli that the participants were viewing (*i.e.*, either risky or not risky images), sent via a parallel port by the PC used to run the task. The signal was recorded with a 1000 Hz sampling frequency and then downsampled at a 30 Hz frequency to be easily exported and stored in a .txt file.

The .txt files were then imported in Matlab and analyzed using the Ledalab toolbox (version 3.4.9; Benedek & Kaernbach, 2010a, 2010b). First, we visually inspected the signal, looking for artifacts (Boucsein, 2012). No artifacts were found across all the recordings. Then, an adaptive smoothing procedure (using a moving Gaussian window with a maximum width of 0.0003 samples as automatically implemented in Ledalab) was applied to the signal, in order to reduce any slight residual noise. After this, we decomposed the signal into its tonic and phasic components using the continuous deconvolution analysis (CDA; Benedek & Kaernbach, 2010a, 2010b) to identify, across the phasic component, the skin conductance responses (SCRs) elicited by the stimuli (that is, the driving images) presented in phase B. A SCR was defined as an increase in the phasic part of the signal of at least 0.05 μS in a time-window between 1 and 5 s after the onset of the stimulus (*i.e.,* the traffic image), as suggested by Boucsein et al., 2012. The CDA method was used to account for both inter- and intra-individual variability in signal shape, as well as the superimposition, rise, and recovery time of the SCRs (Benedek & Kaernbach, 2010a). We decided to consider in the analyses only the SCRs clearly elicited by the stimuli, for two reasons. The first is that, in line with our theoretical framework (see Bechara et al., 1996, 1999; Bechara & Damasio, 2005), our working hypothesis was that the feedback should increase the physiological activation elicited by the stimuli, and that this increased activation (promoted but not directly elicited by the feedback) should in turn improve the participants' behavior. The second is that the EDA is a so-called "slow" signal and therefore the inclusion of later SCRs in the analysis (*e.g.*, SCRs elicited in a 1–5 s window following feedback in order to study physiological activation as a direct consequence of feedback) may lead to confounding results, due to the overlapping of SCRs previously elicited by the stimulus (*i.e.*, the traffic image) on those elicited by the feedback.

2.4. Procedure

The participants completed the tasks in two separate experimental sessions, in which phase A and B were completed following a between-participants counterbalanced order. Each session lasted approximately 45 min. Before starting with the experimental procedure, each participant read and signed the informed consent form and filled in a demographic data sheet. A representation of the experimental setting (common features of both Study 1 and Study 2) is reported in Fig. 2.

Phase A consisted of two tasks (*i.e.,* Evaluation and Decision) presented in a between-participants counterbalanced order. In both tasks, 120 images showing traffic situations were displayed twice in a randomized order. Each trial (see Fig. 3) included a fixation cross presented with a variable duration (between 500 and 750 ms), followed by the stimulus (*i.e.,* the image showing a traffic situation, either risky or not risky) until the participant responded or up to a maximum of 2000 ms. In the case of the Evaluation task, the participant had to evaluate whether the image was risky or not by using the mouse buttons. They had to press the left mouse button with their left forefinger and the right mouse button with their right forefinger. The correspondence between the right and the left mouse button and type of response was counterbalanced between participants. In the case of the Decision task, the participant had to decide whether to brake or not, always using the mouse buttons as explained in the Evaluation task. Depending on the group—either Feedback or Control—the image was followed by a feedback screen or a blank screen (1700 ms). The total duration of a trial was up to 4450 ms, depending on the variable duration of the fixation cross, and each task lasted no more than 20 min. The feedback screen could display either a neutral message ("You keep the points of your driving license") or a contingent negative feedback ("You lose 1 point from your driving license") in 50% of the cases in which the participant made an error. The participants performed two practice trials before starting with the task. The order of the tasks was balanced between participants and no more than 5 min were scheduled between the end of the first task and the beginning of the second.

Phase B (see Fig. 3) included only the Decision task, but all the participants were presented with two conditions, that is, with or without feedback. In both conditions, 40 images showing traffic situations were displayed in a randomized order. The order of the conditions was counterbalanced between participants. Moreover, we measured the participants' EDA for the entire duration of the task. At the beginning of the experiment, we gently cleaned the area of interest of the left foot with a non-alcoholic disinfectant solution. We then applied the electrodes and, after asking the participant to try not to move the foot for the entire duration of the task, we checked the signal quality by asking him/her for three times to take a deep breath in, hold the breath for three seconds, and to then breathe out. Finally, a 3-min baseline without any external stimulation was recorded. Another baseline was recorded between the two conditions. In order to adequately record the EDA signal, which requires long trial durations to prevent overlaps between the SCRs (Boucsein et al., 2012), we adapted the procedure in terms of the task and trial durations. Each trial included a 10-sec fixation cross, followed by a 2-sec duration image of traffic situations, and, in the case of the condition with feedback, the feedback screen (7 sec), which could include either the neutral message or the negative feedback. The total duration of a trial was 19-sec for the condition with feedback and 12-sec for the condition without feedback. The participants received the same instructions as those given during the Decision task in phase A and performed two practice trials before starting the task.

2.5. Experimental design and analyses

Phase A followed a 2 × 2 × 2 repeated-measures mixed design, with *Group* (Feedback *vs*. Control) as the between-subjects variable, and *Task* (Evaluation *vs*. Decision) and *Risk* (Risky *vs*. Not risky images) as the within-subjects variables. Regarding the behavioral analyses of phase A, for each participant for both risky and not risky images in both tasks and groups, we calculated the mean probability of response (PR) (*i.e.,* the probability to brake [Decision task] or to evaluate an image as risky [Evaluation task]; meant as the proportion between the number of trials in which the participant decided to brake or to evaluate an image as risky, and the total number of trials in each condition) (see also Maldonado et al., 2020; Torres et al., 2017)

Then, according to the experimental design, we conducted a mixed repeated-measures analyses of variance (ANOVA) on the PR with *Group*, *Task*, and *Risk* as the independent variables. The design aimed at identifying behavioral differences in the participants' PR as a function of the feedback, the task, and the type of stimuli.

Concerning phase B, we used a 2 × 2 × 2 repeated-measures design with *Feedback* (Feedback *vs*. No Feedback), *Block* (First *vs*. Second block; the first block included the first 20 trials of each condition, while the second included the last 20 trials) and *Risk* (Risky *vs.* Not risky images) as the within-subjects variables. This time, the participants were not split into groups and all of them faced a condition with feedback and a condition without feedback. This aspect of the procedure was designed to minimize the possible random noise that might have been present in a between-participants design, particularly in a design involving a single type of task (in this case, a Decision task). For the behavioral analyses of phase B, we again calculated the PR for both risky and not risky images in both conditions (with and without feedback). However, in order to better assess the changes in participants' behavior throughout the task, we calculated the mean PR for each participant in each block (*i.e.,* 20 trials) of each condition. We then conducted a repeated-measures ANOVA on the PR with *Feedback*, *Block*, and *Risk* as the independent variables. For the EDA signal, we extracted the amplitude of the first SCR after stimulus (*i.e.,* increase of at least 0.05 μS) in each trial, considering a time window of 1–5 sec after the onset of the stimulus (*i.e.,* risky or not risky image). For each participant we calculated the mean SCR amplitude and the percentage of SCRs (*n* trials with at least one SCR * 100/*n* trials in each condition) for risky and not risky images (in each condition and block). One participant (a female) was discarded from the amplitude analysis, since she never displayed a SCR in any of the conditions. Moreover, given that some features of the EDA signal and SCRs (see Boucsein, 2012; Boucsein et al., 2012, Berntson, Cacioppo, & Tassinary, 2017) can result in an absence of SCRs in certain conditions (*e.g.,* Tagliabue et al., 2019), we substituted the missing values in the amplitude data (27% of the total) for the mean of the condition. We then ran two separate repeated-measures ANOVAs, one for the mean SCR amplitude and another for the percentage of SCRs, with *Feedback*, *Block*, and *Risk* as the independent variables.

Moreover, in order to take into account also possible differences due to the unequal distribution of participants with and without a driving license, we run the same analyses removing the two participants that did not hold a driving license. The results, that are substantially comparable to those reported in the following sections, are included as Supplementary Material.

The α level was set at 0.05. All the statistical analyses were conducted with the IBM SPSS 23 statistical software package. All the post-hoc comparisons were conducted using the Bonferroni correction method.

2.6. Results

2.6.1. Phase A

Table 1 summarizes the results of phase A. *Group* and *Risk* reached significance with $F(1,30) = 9.26$, $p = 0.005$, $\eta_p^2 = 0.24$ and $F(1,30) = 771.38$, $p < 0.001$, $\eta_p^2 = 0.96$, respectively. Specifically, the Feedback group showed a significantly higher PR (Mean [M] = 0.59) than the Control group (M = 0.47) in both tasks, and the PR was significantly higher for risky (M = 0.84) than for not risky images (M = 0.22) in both groups (see Table 1). The *Task* was also significant, $F(1,30) = 4.18$, $p = 0.05$, $\eta_p^2 = 0.12$, with the data indicating a higher PR for the Decision task (M = 0.55) than for the Evaluation task (M = 0.51). Moreover, the *Task* × *Risk* interaction reached significance, $F(1,30) = 6.28$, $p = 0.02$, $\eta_p^2 = 0.17$, with a higher PR on the Decision task than the Evaluation task, but only for risky images (M = 0.88 *vs*. 0.81 for Decision *vs*. Evaluation task respectively, $p = 0.02$; but M = 0.23 and M = 0.22 for Decision *vs*. Evaluation of not risky images). No other sources of variance reached significance.

2.6.1.1. Phase A: Brief discussion. These results replicate previous findings on the PR using an Italian sample (first aim of the present study) (Maldonado et al., 2020; Megías et al., 2015; Megías, Cándido et al., 2018; Megías, Torres et al., 2018; Torres et al., 2017). Taken together, these findings support the idea that both tasks could be controlled by different processes, with the Decision task showing a higher PR and being more prone to the effects of feedback. Thus, the rational-analytic and the experiential-affective system could differentially control each task, with the second system being more strongly related to the Decision task and more prone to the effects of feedback, that is, to an increase in the PR when feedback is given, as previous studies have already shown in terms of reaction times (part of the third aim of the present study) (Maldonado et al., 2020; Megías et al., 2015). This evidence suggests the need for a more in-depth investigation of the processes underlying performance on the Decision task, as we did in phase B.

Table 1
Summary of the results (mean and, in parentheses, standard errors) of Study 1 – phase A.

	Risky		Not risky		
Phase A	Decision	Evaluation	Decision	Evaluation	*Total Group*
Control	0.83 (0.02)	0.72 (0.02)	0.18 (0.04)	0.16 (0.04)	**0.47 (0.03)**
Feedback	0.92 (0.02)	0.89 (0.04)	0.27 (0.04)	0.28 (0.04)	**0.59 (0.03)**
Total	**0.88 (0.01)**	**0.81 (0.03)**	**0.23 (0.03)**	**0.22 (0.03)**	

2.6.2. Phase B

2.6.2.1. Behavioral results. With regard to the behavioral results (Table 2), the *Risk* variable again reached significance, $F(1,31) = 1168.03$, $p < 0.001$, $\eta_p^2 = 0.97$. When faced with risky images, participants showed a significantly higher PR (M = 0.91) than when faced with not risky images (M = 0.14). Moreover, the *Block* × *Risk* interaction reached significance, $F(1,31) = 6.55$, $p = 0.02$, $\eta_p^2 = 0.17$. Post-hoc comparisons revealed significant differences between risky and not risky images in each block, but this difference was greater on the second block (M = 0.16 *vs.* 0.89 for not risky *vs.* risky images respectively in the first block; M = 0.12 *vs.* 0.93 in the second block; all comparisons with $p < 0.001$). The absence of an effect of the *Feedback* factor could be due to a ceiling effect, as the task was very easy due to the type of images used (high-risk) and the task (Decision only), which was repeated twice. To illustrate, the ceiling effect may be due to the features of selected traffic images, that were all classified as either high-risk or low-risk, meaning that their associated risk values (derived from the validation of the dataset, see Section 2.2) were very different and, therefore, that their features may have helped the participants in easily discriminate between them.

2.6.2.2. EDA results: SCR amplitude. The most relevant results come from the EDA analyses (Fig. 4). A significant effect of *Feedback* was found in the case of stimulus-elicited mean SCR amplitude, $F(1,30) = 13.70$, $p = 0.001$, $\eta_p^2 = 0.31$. We observed a significantly higher amplitude of SCRs in the Feedback condition (M = 0.39) in comparison with the No Feedback condition (M = 0.26). Moreover, although the *Feedback* × *Risk* interaction was not significant, $F(1,30) = 3.24$, $p = 0.08$, it seemed to indicate a tendency towards an increased SCRs amplitude in the case of the Feedback condition (M = 0.41) compared with the No Feedback condition (M = 0.24) for not risky images.

For the stimulus-elicited percentages of SCRs, a significant effect was found for the *Block* variable, $F(1,31) = 53.04$, $p < 0.001$, $\eta_p^2 = 0.63$. The percentage of SCRs significantly decreased during the second block of the task, (M = 42.80 *vs.*28.41 in the first and the second block respectively). Finally, although the *Feedback* × *Block* interaction was not significant, $F(1,31) = 3.51$, $p = 0.07$, the data seem to indicate a tendency towards an increased percentage of SCRs in the Feedback condition (M = 46.36%) with respect to the No Feedback condition on the first block (M = 39.23%).

2.6.2.3. Phase B: Brief discussion. The EDA results of phase B indicate that, in the conditions where a contingent negative feedback was delivered, an increase in the participants' psychophysiological activation (in terms of SCR amplitude in response to the stimuli) was observed. This effect was more marked in the first part of the task, as shown by the higher percentage of SCRs in the first block (part of the second aim of the present work). This evidence favors an associative explanation of the effect of feedback, in which an association is formed between stimulus features, behavioral response and penalty during the first part of the task, paralleled by an increase in psychophysiological activation for the stimulus itself, which could be a consequence of somatic marker activation, as seen in the studies using a moped driving simulator (Tagliabue et al., 2017, 2019; Tagliabue & Sarlo, 2015) (part of third aim of the present work). The decrease in skin conductance-related measures in the second part of the task may be due to the fact that a lower level of emotional activation is needed once the safe behavior has already been established.

The main limitation of Study 1 is related to the features of the experimental paradigm used in phase B, (1) namely the difference in the trial duration, (2) the differences in structure between the condition with feedback and that without feedback, which may have masked the behavioral effect of feedback as well as some EDA effects, due to the difficulty in comparing the two conditions, and (3) the probability with which the negative feedback was administered (*i.e.*, 50% of the trials in which the participants made an error), that may lead to a very infrequent administration of the negative feedback in case of low error frequency. Moreover, the EDA results of this study only refer to the decision task, leaving unexplored the psychophysiological mechanisms underlying the evaluation task. Finally, the absence of clear differences in the behavioral findings is most likely due to a ceiling effect resulting from the characteristics of the selected traffic images. Taken together, these considerations led us to design Study 2.

3. Study 2

In Study 2 we aimed to overcome the limitations of Study 1 by improving the experimental paradigm. To this end, we included a set of medium-level risk images (see Fig. 1) in order to avoid a ceiling effect and to strengthen the feedback, also increasing the probability with which the feedback was administrated (see below). Moreover, we designed the task to ensure an identical trial duration and structure in all conditions. Finally, we included both a decision and an evaluation task in order to assess and compare the psycho-physiological mechanisms underlying the effect of feedback in both tasks.

Table 2
Summary of the behavioral results (mean and, in parentheses, standard errors) of Study 1 – phase B.

	Risky		Not risky	
Phase B	Block 1	Block 2	Block 1	Block 2
No Feedback	0.90 (0.03)	0.93 (0.02)	0.12 (0.02)	0.13 (0.02)
Feedback	0.88 (0.02)	0.93 (0.02)	0.20 (0.04)	0.12 (0.03)
Total	**0.89 (0.02)**	**0.93 (0.02)**	**0.16 (0.03)**	**0.12 (0.02)**

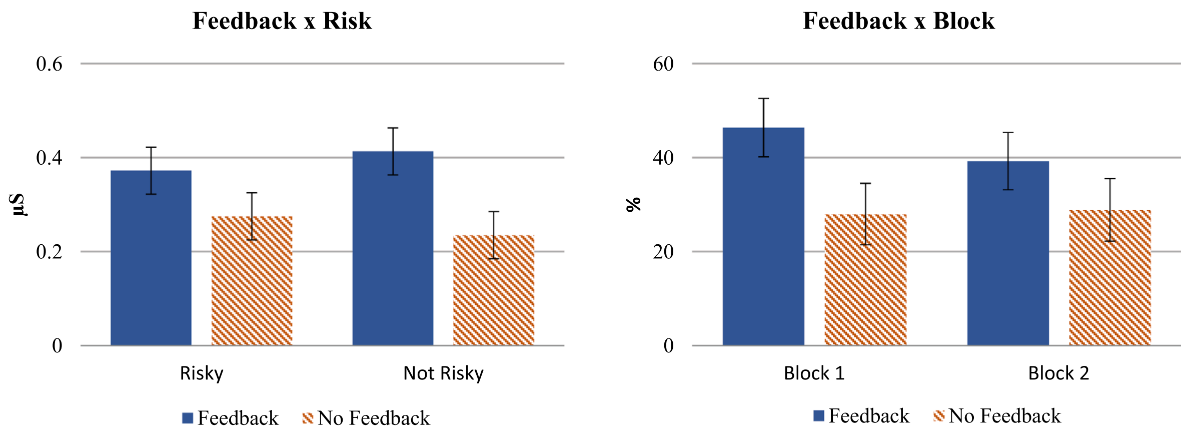

Fig. 4. Psychophysiological results (SCR amplitude, left panel; percentage of SCRs, right panel) of Study 1 – Phase B. Vertical bars represent standard errors.

3.1. Participants

We used again a convenience sampling method to recruit forty-four Spanish students from the University of Granada that agreed to take part in the study (mean age ± SD = 19.93 ± 2.42 years; age range: 18–30 years; 35 females). Twenty-two participants were assigned to the Feedback group (mean age ± SD = 20.45 ± 3.04 years; age range: 18–30 years; 18 females; 9 holding a valid car driving license) whereas the other to the Control group (mean age ± SD = 19.41 ± 1.50 years; age range: 18–24 years; 17 females; 9 holding a valid car driving license). All the participants had normal or corrected to normal vision, did not report any neurological disorder or any use of medications. The participants did not receive any compensation for their participation in the study. All the participants gave informed consent and were assured of their rights according to the Code of Ethics of the World Medical Association (Declaration of Helsinki – WMA, 2008), under which suggestions the study was conducted.

3.2. Stimuli

Eighty images showing actual first-person traffic situations were used in the study (forty for each task, as in phase B of Study 1). The images were selected from the same extensive database of traffic situation pictures used in Study 1 (see Megías, Torres et al., 2018). A key difference with respect to phase B of the previous study was that forty of the selected pictures showed a medium risk level (average risk = 3.65 [0: no risk; 7: high-risk]) and the other forty showed no risk (average risk = 1.98). As in Study 1, all the selected images met a set of statistical criteria to reduce interpersonal variability (see Section 2.2). In this way we aimed at enhancing the effect of feedback in comparison with Study 1. The stimuli were displayed through E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA – version 2.0.10) using a Windows 7 system PC, with a 14-inch monitor and a resolution set to 1024 × 768, placed at a distance of ≈ 75 cm from the participant.

3.3. EDA recording and processing

EDA was recorded throughout the entire duration of the task employing two Ag/AgCl electrodes (1 cm diameter), filled with K-Y lubricating jelly. The electrodes were again placed on the participants' left foot, as in Study 1 (see Section 2.3) (Boucsein et al., 2012).

An AC galvanic skin response (ML116 GSR Amp) amplifier, associated with a PowerLab 8/30 series recording unit (ADInstruments Ltd., Dunedin, NZ) providing a 0.2-V constant voltage across electrodes with a 10 Hz low-pass filter was used to record the participants' EDA signal. The amplifier was connected to a Windows 7 system PC running the LabChart software (ADInstruments Ltd., Dunedin, NZ – versionv5.2.1) for the recording of the EDA. The system recorded and displayed on a dedicated monitor, not visible to the participants, the signal and the markers corresponding to the different stimuli the participants were viewing (either risky or not risky images), sent through a parallel port by the PC on which the task was run. The signal was recorded, exported, and stored in a .txt file with a 40 Hz sampling frequency. The .txt files were then imported in Matlab and analyzed by means of the Ledalab toolbox (version 3.4.9; Benedek & Kaernbach, 2010a, 2010b) following the same procedure described in Study 1 (see Section 2.3). Again, we decided to focus on the SCRs elicited by the stimulus. The rationale for this decision is described in Section 2.3.

3.4. Procedure

The participants were invited to take part in a single experimental session that included both tasks (*i.e.,* Evaluation and Decision) (see Fig. 2 for a representation of the general experimental setting). The session lasted approximately 45 min. Before starting the experimental procedure, each participant read and signed the informed consent form and filled in a demographic data sheet. We then applied the electrodes with a non-alcoholic disinfectant solution. An initial 3-min baseline without any external stimulation was recorded. A second baseline was recorded between the two tasks, administered in a between-participants counterbalanced order.

In both tasks, we displayed 40 images depicting traffic situations. On each trial (see Fig. 5) a fixation cross (500 ms) appeared, followed by the stimulus (*i.e.,* the image showing a traffic situation, either risky or not risky) until the participant responded, or a maximum of 2000 ms had elapsed. As in phase A of Study 1, in the case of the Evaluation task, the participant had to evaluate whether the image was risky or not by pressing the mouse buttons. The correspondence between the right and the left mouse button and type of response was counterbalanced across participants. In the case of the Decision task, the participant had to decide whether to brake or not, always using the mouse buttons, as explained. Depending on the group, either Feedback or Control, the image was followed by a feedback screen or a neutral message screen (2000 ms). Finally, a blank screen was shown in order to allow the measurement of late SCRs. Another blank screen with a variable duration was displayed immediately after the participant's response in order to precisely cover the 2000 ms time window before the appearance of the feedback or the neutral message. The total duration of each trial was fixed at 6500 ms. The feedback screen could display either a neutral message ("You keep the points of your driving license") or a contingent negative feedback ("You lose 1 point from your driving license"). The negative feedback was delivered in 75% of the cases in which the participant, when presented with a risky image, did not decide to brake or did not evaluate the scenario as risky. In addition, the negative feedback was always delivered after making the first two errors. The participants performed two practice trials before starting the task. We decided to increase the percentage of feedback administration and to always administer it after the first two errors in order to maximize the feedback effect by ensuring that, unlike in Study 1, in the case of error, each participant received a minimum number of negative feedbacks. The order of the tasks was balanced between participants and no more than 5 min were scheduled between the end of the first task and the beginning of the second.

3.5. Experimental design and analyses

Treasuring the results and the limits of the previous experiments, we designed a 2 × 2 × 2 repeated-measures mixed design with *Group* (Feedback *vs.* Control) as the between-subjects independent variable and *Task* (Evaluation *vs.* Decision) and *Block* (First *vs.* Second block) as the within-subjects independent variables. Moreover, throughout both tasks, we recorded the participants' EDA (see Section 3.4). The order of the two tasks was counterbalanced between participants to prevent sequence effects. The aim of this design was to identify behavioral (PR) and physiological (EDA) differences as a function of the feedback, the task, and, given its importance for physiological effects, the block (see Study 1- phase B).

As in Study 1, we calculated the mean PR for each participant for both risky and not risky images on each task. But in this case, we also included blocks, having reduced the number of trials in comparison with phase A of Study 1, and to look for a block effect similar to the one found in phase B. We then conducted two separate repeated-measures ANOVAs on the PR of risky and not risky images with *Group*, *Task,* and *Block* as the independent variables.

For the EDA signal, as in Study 1 we extracted the amplitude of the first SCR after the stimulus presentation and then calculated the mean SCR amplitude and the mean percentage of SCRs on each trial for each participant, for both risky and not risky images (in each task and block). The SCR amplitude data of two participants (two females) were discarded because they did not have any SCR in any of the conditions. Missing values of amplitude were again substituted with the mean of the condition (30% of the total). We then ran two

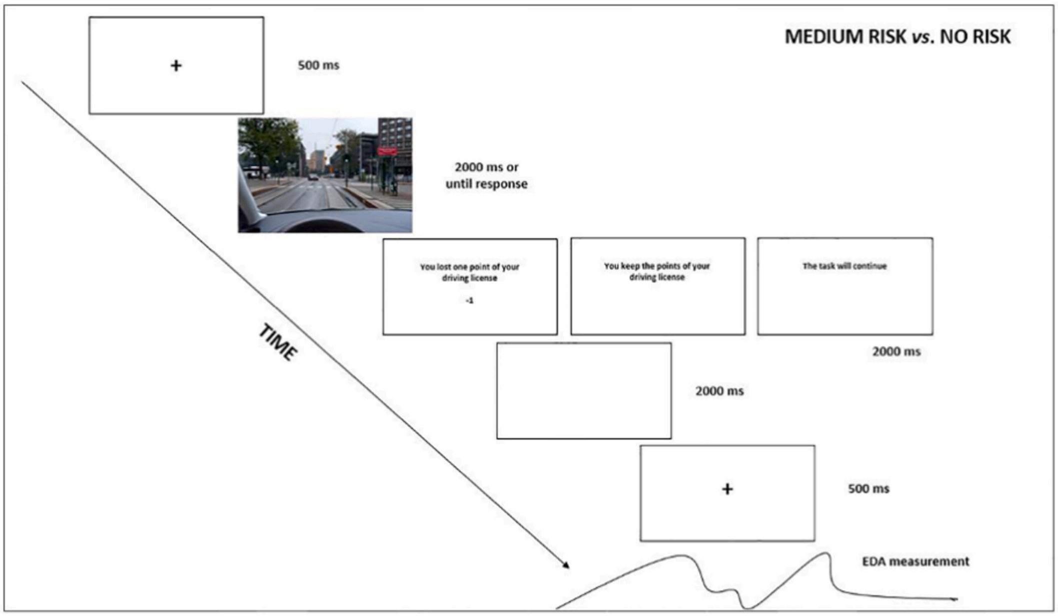

Fig. 5. Schematic representation of a trial for the tasks included in Study 2. Throughout both tasks, we measured the participants' EDA (represented by a curved line resembling the typical shape of the signal at the bottom of the figure).

separate repeated-measures ANOVAs for each dependent variable (*i.e.,* mean SCR amplitude and percentage of SCRs), one for risky images and one for the not risky images, with *Group*, *Task*, and *Block* as the independent variables.

Finally, in order to assess any potential effects linked to the presented stimuli, we ran two further separate repeated-measures ANOVAs on each dependent variable, one for the Evaluation task and another for the Decision task, with *Group*, *Block*, and *Risk* (Risky *vs.* Not risky images; within-participants) as the independent variables.

The α level was set at 0.05. All the statistical analyses were conducted with the IBM SPSS 23 statistical software package. All post-hoc comparisons were run with the Bonferroni correction method.

3.6. Results

3.6.1. Behavioral results

3.6.1.1. Risky images. In the case of risky images, both *Group* and *Task* reached significance, $F(1,42) = 14.33$, $p < 0.001$, $\eta_p^2 = 0.25$ and $F(1,30) = 10.92$, $p < 0.001$, $\eta_p^2 = 0.21$ respectively, similarly to phase A of Study 1 (see Table 3). As in the previous study, for risky images the group that received feedback showed a significantly higher PR than the control group, with the PR also being higher for Decision task than Evaluation task.

It is important to note how the main difference between the two studies was the higher level of risk shown by the risky images used in the previous study, which could explain the slightly better performance shown by participants in Study 1. Moreover, the *Group* × *Block* interaction reached significance, $F(1,42) = 6.00$, $p = 0.02$, $\eta_p^2 = 0.13$. Post-hoc comparisons revealed a significant difference between groups in the second block, with the Control group again showing a significantly lower PR than the Feedback group (M = 0.68 *vs.* 0.86, $p < 0.001$) and with a significant increase in the PR of the Feedback group from the first to the second block (M = 0.79 *vs.* 0.86, $p = 0.02$). No differences were found between the two groups on the first block, and no changes were evident in the PR of the Control group from the first to the second block.

3.6.1.2. Not risky images. For the not risky images, only the effect of *Group* was significant, $F(1,42) = 4.95$, $p = 0.03$, $\eta_p^2 = 0.11$, since participants in the Feedback group showed a significantly higher mean PR than those in the Control group, as also observed in Study 1 (see Table 1).

3.6.2. EDA results

3.6.2.1. Risky images. Concerning the risky images (Fig. 5), analysis of the mean stimulus-elicited SCR amplitudes revealed an almost marginally significant effect of both *Block* and a *Group* × *Block* interaction, $F(1,40) = 3.84$, $p = 0.06$ and $F(1,40) = 3.24$, $p = 0.08$, respectively. The data seem to indicate a tendency towards a higher SCR amplitude for the Feedback group (M = 0.27) than the Control group (M = 0.17) on the second block, with a decrease in SCR amplitude for participants in the Control group from the first (M = 0.27) to the second block (M = 0.17).

Analysis of the percentage of stimulus-elicited SCRs revealed a significant effect of both *Block* and the interaction *Group* × *Block*, $F(1,42) = 70.03$, $p < 0.001$, $\eta_p^2 = 0.63$, and $F(1,42) = 4.57$, $p = 0.04$, $\eta_p^2 = 0.10$, respectively. Post-hoc comparisons showed a significant decrease in the percentage of SCRs from the first to the second block for both the Control (M = 34.64% *vs.* 19.29%; $p < 0.001$) and the Feedback group (M = 47.54% *vs.* 21.66%; $p < 0.001$) (Fig. 6).

3.6.2.2. Not risky images. Concerning the not risky images, analysis of the stimulus-elicited mean SCR amplitudes revealed a significant effect of both *Task* and *Block*, $F(1,40) = 4.45$, $p < 0.001$, $\eta_p^2 = 0.10$ and $F(1,40) = 9.28$, $p < 0.001$, $\eta_p^2 = 0.19$, respectively. Higher SCR amplitudes were found in the Decision task (M = 0.25) in comparison with the Evaluation task (M = 0.21) and, overall, in the first block the participants showed higher SCR amplitudes (M = 0.26) than in the second block (M = 0.20). Moreover, the *Group* × *Task* interaction was marginally significant with $F(1,40) = 4.07$, $p = 0.05$. The data indicate a tendency towards an increased SCR amplitude in the Feedback group for the Decision task (M = 0.28) compared with the Evaluation task (M = 0.21). In the case of the stimulus-elicited percentage of SCRs, the *Block* is the only source of variance that reached significance, $F(1,42) = 51.13$, $p < 0.001$, $\eta_p^2 = 0.55$: a significant decrease in the percentage of SCRs was evident from the first to the second block (M = 40.70% *vs.* 19.63% respectively).

3.6.2.3. Evaluation task. Finally, we separately examined the EDA data from the two tasks. In the case of stimulus-elicited SCR

Table 3
Summary of the behavioral results (mean, and in parentheses, standard errors) of Study 2.

	Risky		Not risky		
	Decision	Evaluation	Decision	Evaluation	*Total Group*
Control	0.76 (0.02)	0.63 (0.04)	0.08 (0.04)	0.13 (0.03)	**0.40 (0.03)**
Feedback	0.85 (0.02)	0.81 (0.04)	0.19 (0.03)	0.21 (0.03)	**0.51 (0.03)**
Total	**0.81 (0.02)**	**0.72 (0.04)**	**0.14 (0.03)**	**0.17 (0.03)**	

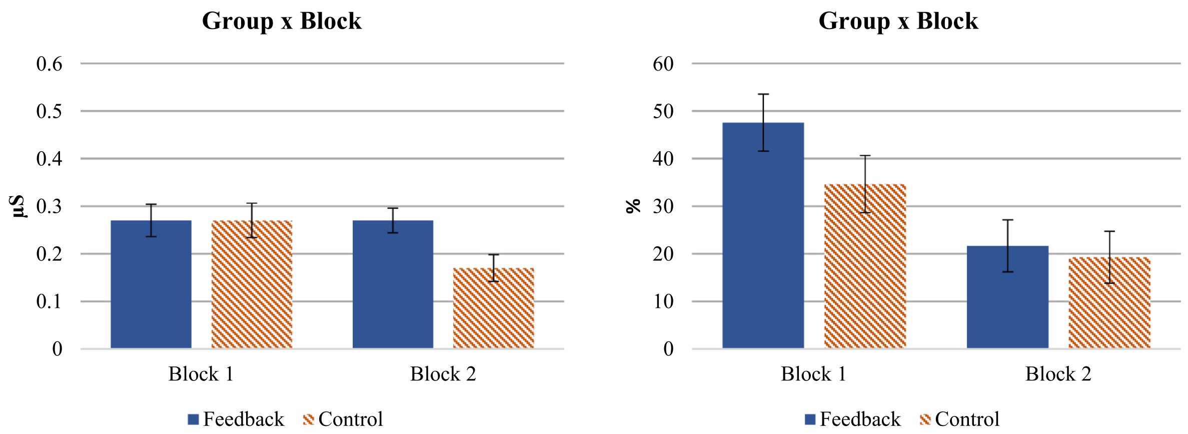

Fig. 6. Results for SCR amplitude (left panel) and percentage of SCRs (right panel) for risky images in Study 2. The results show a significant decrease in the percentage of SCRs between the two blocks in both groups. However, on the second block, the Feedback group showed a higher SCR amplitude than the Control group. Vertical bars represent standard errors.

amplitude for the Evaluation task, the *Risk* variable reached significance, $F(1,40) = 4.69$, $p = 0.04$, $\eta_p^2 = 0.11$. The participants showed significantly higher SCR amplitude when facing risky images (M = 0.25) in comparison with not risky images (M = 0.21). Notably, although both the *Block* variable and the *Group* × *Risk* interaction were not significant, $F(1,40) = 3.27$, $p = 0.08$ and $F(1,40) = 3.25$, $p = 0.08$, respectively, the data tendencies seem to indicate a general increase in the SCR amplitude in the second block (M = 0.26) and a higher SCR amplitude in the Feedback group (M = 0.29) than the Control group (M = 0.22), but only in the case of risky images.

The analysis of the stimulus-elicited percentage of SCRs revealed a significant effect of the *Block* variable, $F(1,42) = 30.36$, $p < 0.001$, $\eta_p^2 = 0.42$. The percentage of SCRs was higher in the first block of the Evaluation task (M = 38.63%) than in the second block (M = 19.87%). Moreover, the *Group* × *Block* × *Risk* interaction reached significance, $F(1,42) = 5.55$, $p = 0.02$, $\eta_p^2 = 0.12$. Post-hoc comparisons revealed that the percentage of SCRs significantly decreased between blocks, independently of the type of images (either risky or not risky; all comparisons with $p < 0.05$). However, the Feedback group showed a tendency toward a less sharp decrease compared with the Control group, particularly in the case of not risky images (M = 41.3% *vs.* 23.7% for the Feedback group in the two blocks, indicating a decrease of 17.6%; M = 36.7% *vs.* 13.9% for the Control group in the two blocks, indicating a decrease of 22.8%).

3.6.2.4. Decision task. In the case of the Decision task, the analysis of the mean stimulus-elicited SCR amplitudes revealed a significant effect of *Block*, $F(1,40) = 9.91$, $p = 0.003$, $\eta_p^2 = 0.20$. As in the Evaluation task, the SCR amplitude was significantly higher in the first block (M = 0.27) than in the second block (M = 0.21). Moreover, although the *Group* × *Block* interaction was not significant, $F(1,42) = 2.95$, $p = 0.09$, the trends suggested by the data seem to indicate that, on the second block of the task, participants in the Feedback group tended to show a higher SCR amplitude (M = 0.25) than those in the Control group (M = 0.17) (Fig. 7).

Analysis of the stimulus-elicited percentage of SCRs revealed only a significant effect of *Block*, $F(1,42) = 51.17$, $p < 0.001$, $\eta_p^2 = 0.55$. A significantly higher percentage of SCRs was found in the first block (M = 43.16%) than in the second block (M = 20.23%).

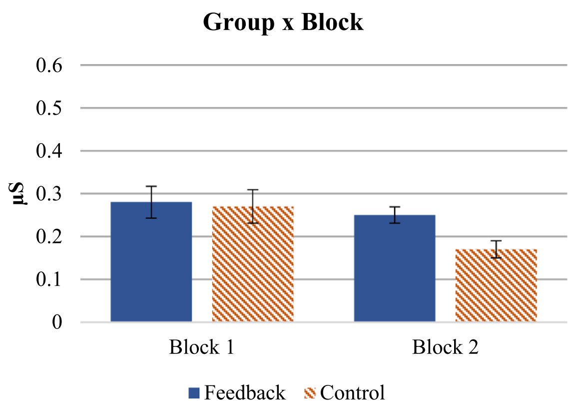

Fig. 7. Mean SCR amplitudes on the Decision task in Study 2. SCR amplitude tends to be lower for the Control group in the second block. Vertical bars represent standard errors.

3.6.3. Study 2: Brief discussion

Overall, the behavioral results replicate those of Study 1 (Phase A), regarding the effect of feedback and task factors. Thus, both effects appear to be independent of the cultural features of the sample studied (either Spanish or Italian – first aim of the present work), and confirm the assumption that two processing systems (*i.e.*, controlled and automatic) differentially control each task (Maldonado et al., 2020; Megías et al., 2015). Moreover, the differential modulation of blocks as a function of feedback in the PR allows for a better understanding of the psychophysiological mechanisms paralleling the effects observed in the previous study, which we have attempted to furtherly analyze using the EDA results of this experiment.

Indeed, in Study 2 we aimed to overcome the limitations of Study 1 in order to assess the psychophysiological correlates of feedback in shaping participants' responses in both Decision and Evaluation tasks, whilst also exploring in further depth the role of task moment (*i.e.*, block). If the behavioral results revealed how participants that received contingent negative feedback showed a higher number of safer responses from the first to the second block when compared with the Control Group, this effect appears to be also confirmed by the EDA results (second aim of the present work). In particular, it appears that in the case of risky situations, there is a higher stimulus-elicited SCR amplitude and percentage of SCRs in the first part of the task compared with the second part; however, this difference is more pronounced for the Feedback group in the case of percentage SCRs. As the task proceeds and the associations are formed (see Study 1), the SCR percentage is reduced in both groups, but, interestingly, the SCR amplitude does not decrease in the Feedback group.

This result can be interpreted through an associative view of how feedback exerts an effect on performance (Feng & Donmez, 2013; Torres et al., 2017). Feedback increases psychophysiological activation for the presented stimuli at the beginning of the task and then, after the association between stimulus features and the correct response has been formed, such activation is reduced due to the fact that a lower emotional impact is needed to continue exhibiting safe behaviors. On the other hand, the persistence of the effect on the SCR amplitude could be the result of an increase in the sensitivity of participants to risky images (third aim of the present work).

In fact, with regard to the not risky images, we found a significant difference between the two tasks in terms of SCR amplitude, with higher values in the Decision task, which further supports the notion that the two tasks rely on different processes. It should be noted that the effect of *Block* was significant in the majority of the analyses for both dependent variables, suggesting that an increase in activation in the first part of the tasks seems to be necessary for performance. This activation is further increased when the feedback is present, apparently leading to an overall safer performance. However, once the correct response has been learned, the activation seems to decrease, although this decrease is overall less evident in the Feedback group, which continues to benefit from the physiological activation promoted by the use of feedback that also produces safer responses in the second block (*i.e.*, the increase in PR for the risky images, independently of the task). This perhaps constitutes the most important source of evidence in favor of an associative explanation of the effects of feedback, as we will see in the next section (second and third aim of the work).

4. General discussion

Consistently with the main dual-process models of decision-making (*e.g.*, Slovic et al., 2007), we hypothesized the presence of two systems, that is, an experiential-affective and a rational-analytic system, whose interplay regulates the decision-making process in risky contexts such as driving. In our paradigm, the Decision task (*i.e.*, deciding whether or not to brake when faced with a set of images representing either risky or not risky traffic situations) is assumed to rely primarily on the experiential affective system, whereas the Evaluation task (*i.e.*, evaluating each image as risky or not) is thought to depend on a process that is more strongly associated with the rational-analytic system (Megías et al., 2011, 2015). In the past, these tasks have proved to be differentially affected by the use of contingent negative feedback (Maldonado et al., 2020), whilst also related to specific electrophysiological correlates and brain activation patterns (Megías et al., 2015). However, to the best of our knowledge, no studies have attempted to assess the psychophysiological mechanisms (*e.g.*, EDA) paralleling the delivery of a contingent negative feedback in risky contexts, although a body of literature suggests that these mechanisms could play a key role in explaining the effect itself (see Bechara et al., 1996, 1999; Bechara & Damasio, 2005; Damasio, 1994). Therefore, starting from the evidence that contingent negative feedback influences risk perception and decision-making during driving, as measured through an Evaluation and Decision task respectively (*e.g.*, Megías et al., 2015; Megías, Cándido et al., 2018; Megías, Torres et al., 2018; Torres et al., 2017), we overall aimed to assess the psychophysiological correlates of the role of feedback in shaping driving behavior.

We conducted two studies. Study 1 aimed to replicate, in an Italian sample, the previous behavioral results obtained with the use of both tasks in various Spanish samples (Megías et al., 2015; Megías, Cándido et al., 2018; Megías, Torres et al., 2018; Torres et al., 2017). Moreover, this study represented the first attempt to empirically explore the possible psychophysiological correlates of the feedback. The data showed a significant increase in the PR (that is, in the safer responses frequency) for both the Decision and the Evaluation task when contingent negative feedback was administered. Moreover, the differences found in the PR between the two tasks, with a higher PR for the Decision task in risky situations, are in line with the postulates that Decision and the Evaluation tasks can be guided by different processes. According to previous studies (Maldonado et al., 2016; Torres et al., 2017), and following the dual-process models of decision making (Kahneman, 2011; Reyna, 2004; Slovic et al., 2004), the Decision task would be mainly controlled by the experiential-affective system, which is more automatic and affective-driven than the rational-analytic system (Megías et al., 2015; Slovic et al., 2007). These characteristics would allow the experiential-affective system, compared with the rational-analytic system, to be more sensitive and prone to safer reactions produced by the need to act urgently (*e.g.*, braking), thus resulting in an increased PR toward the safe response option under risky circumstances. However, the behavioral results of phase B in Study 1 did not reveal any clear moderation of the feedback factor, which was probably the consequence of a ceiling effect due to a sort of over-simplification of the task (only the decision task) administered twice (*i.e.*, presence or absence of feedback).

Moreover, the EDA results of Study 1 pointed to a moderation of the participants' psychophysiological activation, which could possibly underpin the development of the behavioral effects. Indeed, when negative contingent feedback was administered, an increase in psychophysiological activation (in terms of stimulus-elicited SCR amplitude) was recorded, an effect that was more marked in the first part of the task, as shown by the higher percentage of SCRs in the first block. Thus, the effects shown in the Decision task could have been reflected by an increase in physiological activation, which was moderated by the moment of the task. This result supports an associative explanation of the effect of feedback: the increase in physiological activation at the beginning of the task would reflect the formation of associations between the stimulus features, behavioral response, and penalty, leading to an increase in the PR, as shown by an increased physiological activation in response to the stimuli, which should be promoted by feedback. In the second part of the task, once the learning has been completed, this activation would be reduced. Indeed, an attentional explanation of the effect of feedback, as hypothesized in other studies (*e.g.*, Donmez et al., 2007; Torres et al., 2017), should instead lead to a constant physiological activation or, on the other hand, to an increase in the psychophysiological activation in the second part of the task.

However, interpretation of the EDA findings reported in Study 1 are constrained by four main limitations, namely the selection of the images (see above), the inclusion in the experimental paradigm of only the Decision task, the difficulty in comparing the two conditions (Feedback and No Feedback) due to the differences in the length of the trials, and the probability of administration of the negative feedback (50% of times in case of errors) that, in case of low error frequency, may result in a too infrequent administration of the negative feedback. These limitations prevented us from drawing any firm conclusions regarding the psychophysiological correlates of the behavioral results. Thus, we improved our paradigm in order to overcome these limitations and to adequately assess the differences between the Decision and Evaluation task, as moderated by the moment of the task and the type of image (Study 2).

The behavioral results of Study 2 were in line with previously reported findings (*e.g.*, Megías, Cándido et al., 2018; Torres et al., 2017) and with the results of Study 1, although in this study the data were collected from a sample of Spanish participants, thus suggesting again that, apparently, the feedback moderation is independent from cultural specificities. Indeed, the behavioral moderations linked to feedback presence were confirmed, and the data also replicated the previously reported difference in the PR between the Decision and the Evaluation task in risky situations. Moreover, the behavioral findings of Study 2 confirmed that introducing the EDA recording procedures throughout the task did not have an impact on the participants' behavior. With respect to the EDA results, we confirmed that the physiological activation is clearly moderated by Block, findings that support the associative explanation discussed previously. On the other hand, the differences observed in the psychophysiological activation between the two tasks (Decision and Evaluation) are in line with the assumption that there are differences in the way in which the information is processed in each task as a function of the involvement of the experiential-affective system (Megías et al., 2011, 2015; Slovic et al., 2007).

4.1. General limitations and future steps

In addition to the already mentioned limitations of Study 1, the present work has other limitations that need to be considered. First, although the number of trials on each task and the use of mixed designs helped to minimize the effects of inter-individual differences, the sample size in both studies should be extended in future studies. Second, the characteristics of the samples (university students) and the use of convenience recruitment procedures may have caused some sort of sampling bias that may represent a limitation for the generalization of our results to different populations. Third, the results regarding the effect of Block (*i.e.*, a decrease in psychophysiological activation in the second part of the tasks) could also be interpreted in terms of the effect of habituation of the EDA signal. However, this could not explain the fact that the effect seems to be modulated, at least to a certain degree, by feedback. In any case, the disentanglement of these two aspects may be an interesting challenge for future studies. Fourth, the absence of mediation analyses (that will be included in the next steps to be performed in this research line) prevents us from drawing conclusions regarding the possible causality of the relations between feedback, psychophysiological activation, and behavioral results. Still, the interaction/moderation effects reported, as well as the methodology employed for the psychophysiological assessment (see below), are enough to talk about psychophysiological correlates of the observed behavioral effects. In psychophysiological studies this is usually enough to draw at least initial conclusions regarding the relation between behavioral results and underpinning psychophysiological mechanisms.

Fifth, one may wonder why the EDA analysis was focused on the SCRs elicited by the stimuli (the road scenario) and not those produced by the feedback itself. As already partially explained (see Section 2.3), this decision was prompted by both theoretical and practical considerations. First, we postulated that the feedback should promote physiological activation, as shown by an increase in amplitude and percentage of stimulus-elicited SCRs, in line with evidence reported in the literature (see Bechara et al., 1996, 1999; Bechara & Damasio, 2005). Note that this activation, which is assumed to drive consequent behavior, should be promoted by feedback, meaning that the conditions/groups in which feedback is delivered should show greater activation in response to the stimulus (not the feedback itself). The second reason is that attempting the analysis of the SCRs elicited after the onset of feedback would have carried the risk of also including the SCRs previously produced by the stimuli, resulting in unclear and even biased results (see also Section 2.3).

Sixth, another aspect that deserves consideration, is that, from a behavioral perspective, in the present study no reaction time data were recorded, which potentially prevents us from drawing firm conclusions regarding the dissociation between the two systems. We did not include reaction time measures for two main reasons: the first is that, given the widely documented evidence regarding differences in reaction times between the two tasks (*i.e.*, faster reaction times in the decision task, modulated by feedback when delivered; see Megías et al., 2015; Megías, Cándido et al., 2018; Megías, Torres et al., 2018; Torres et al., 2017), a further replication of these results seemed unnecessary. The second reason is that historically, PR results tend to be unclear and our findings could shed light on this issue. Moreover, as already discussed, the differences that emerged between the Decision and the Evaluation task in terms of physiological activation support the dissociation between the two processing systems, which adds to our existing knowledge of these mechanisms.

Finally, one may also wonder whether the use of static traffic images may be reasonably justified to assess risk perception and how the present results, obtained in a laboratory context and with tasks very different from driving a vehicle, can transfer to real world. Without any doubt, considering that driving a real vehicle requires an overall different cognitive and physical engagement from that required by the tasks we used in our work, it is not possible to ensure a straightforward translation of our results to the real word. This is a common issue in the road-safety investigation field, that even the use of a driving simulator cannot totally overcome.

However, we employed well-known risk perception assessment methods, that include, beside the use of simulators (*e.g.*, Megías et al., 2011; Megías et al., 2017; Tagliabue & Sarlo, 2015; Tagliabue et al., 2017, 2019), also the use of videos reproducing traffic scenarios (*e.g.*, Chirles, Ehsani, Kinnear, & Seymour, 2021; Kinnear et al., 2013), or even the use of simple traffic images. Regarding traffic images, previous works validated in many different settings the images employed in the present study (*e.g.*, by assessing how the effects of a PC-based risk perception training with these images can translate to the behavior on a driving simulator; Megías et al., 2017). Moreover, it is worth noting that, for instance, the so-called Hazard Perception Test (HPT), in most of its well-known versions (*e.g.*, Scialfa et al., 2012; Tūskė, Šeibokaitė, Endriulaitienė, & Lehtonen, 2019; Wetton et al., 2010), employs traffic images to test and train drivers' risk perception, even, for some countries, in legal settings (*e.g.*, for obtaining or renewing the driving license). Notably, no significant differences were found when comparing versions employing static or dynamic images of the HPT in the ability to predict drivers' self-report driving behavior (Scialfa, Borkenhagen, Lyon, & Deschênes, 2013).

Therefore, although we cannot be sure that the identified drivers' responses would be exactly the same in the real world, still the use of a well-known and validated approach (traffic images) allows us to be reasonably confident about the reliability of the data. Moreover, considering that, as it appears from our results, even simple traffic images can elicit psychophysiological activation, that is thought to trigger and parallel the drivers' responses, the same effect, or even stronger, should be evident when immersed in the real world. This activation should, in turn, affect drivers' behavioral responses, also in the real on-road situations. Future studies should focus on the assessment of the transfer of these results also to driving simulators and, even better, real vehicles.

5. Conclusions

Overall, the present work represents the first attempt to assess the psychophysiological correlates, in terms of EDA, of the contingent negative feedback administered to shape participants' risk perception and decision-making in a driving context. Our results have allowed us to draw two main conclusions.

First, the behavioral results from both studies indicate that the PR seems to be moderated by feedback, that is, it increases when the feedback is delivered, leading to safer responses when the participants are faced with a risky situation, an effect that depends on the type of task (Decision or Evaluation). This result is consistent with those of previous studies and highlights the differences between the two tasks, supporting the predictions made by the dual-process theories (*e.g.*, Slovic et al., 2007). Moreover, this finding appears to be independent of the cultural background of the sample, since it has been found in both Italian and Spanish participants. This evidence, that, to the best of our knowledge, has never been previously reported, highlights how, independently from the cultural features that may affect the drivers' risk perception (see Di Stasi et al., 2020), the administration of a contingent negative feedback and its physiological correlates seem to act in the same way for everyone.

Second, the EDA results suggest that the role of feedback in physiological activation, as measured through the stimulus-elicited amplitude and percentage of SCRs, is mainly expressed in the first part of the task, both in the Decision and Evaluation task. This can be taken as evidence to support an associative explanation of the role of feedback in shaping behavior, rather than an attentional account (Torres et al., 2017).

Taken together, the present results are of theoretical relevance to the study of feedback action and of its psychophysiological correlates. From a practical perspective, our findings make a crucial contribution towards understanding risk perception and its role for driving behavior. For instance, they could inform the development of specific, time-effective training and interventions aimed at reducing the rates of driving accidents through the improvement of risk perception and decision-making processes. Moreover, the future possibility to explore the causal relations between feedback, psychophysiological response, and behavior, *via* mediation analyses, may lead to the development of detailed risk perception assessment protocols, to be implemented also in driving license obtainment or renewal contexts. Finally, if successfully applied to driving simulator contexts or to the real world, they may become of crucial importance in the autonomous driving field, giving the chance to develop tailored systems based on the sensed drivers' risk perception, able to safely regulate the transitions between different autonomous driving levels, which is currently one of the main obstacles to the development of this technology.

6. Authors' contribution

EG conducted data collection, statistical and signal analysis, and manuscript writing. MT supervised data collection, statistical and signal analysis, and contributed to results discussion, manuscript writing, and revision. AMR contributed to experiment design, manuscript writing, and revision. AM supervised the experimental designs, data collection and statistical analysis, and contributed to results discussion, manuscript writing, and revision. All authors contributed to research planning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Dr. Miguel Angel Torres for the help with the data collection, and Federico Marafin for graphic support. This work was funded by the Regional Ministry of Economy and Knowledge (Junta de Andalucía, Spain project: UMA18-FEDERJA-137) grant to AMR and carried out within the scope of the project "use-inspired basic research", for which the Department of General Psychology of the University of Padua has been recognized as "Dipartimento di Eccellenza" by the Ministry of University and Research (Italy).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2021.10.001.

References

Aidman, E., Chadunow, C., Johnson, K., & Reece, J. (2015). Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. *Accident Analysis & Prevention, 81*, 8–13.
Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*(13), 5473–5481.
Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex, 6*(2), 215–225.
Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52*(2), 336–372.
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.
Berntson, G. G., Cacioppo, J. T., & Tassinary, L. G. (2017). *Handbook of Psychophysiology*. Cambridge: University Press.
Boucsein, W. (2012). *Electrodermal activity*. New York: 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.
Chirles, T. J., Ehsani, J. P., Kinnear, N., & Seymour, K. E. (2021). Skin conductance responses of learner and licensed drivers during a hazard perception task. *Frontiers in Psychology, 12*.
Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York: Avon.
Di Stasi, L. L., Diaz-Piedra, C., Morales, J. M., Kurapov, A., Tagliabue, M., Bjärtå, A., … Catena, A. (2020). A cross-cultural comparison of visual search strategies and response times in road hazard perception testing. *Accident Analysis & Prevention, 148*, Article 105785.
Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention, 39*(3), 581–590.
Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*(8), 709–724.
Farah, H., Musicant, O., Shimshoni, Y., Toledo, T., Grimberg, E., Omer, H., & Lotan, T. (2014). Can providing feedback on driving behavior and training on parental vigilant care affect male teen drivers and their parents? *Accident Analysis & Prevention, 69*, 62–70.
Feng, J., & Donmez, B. (2013). Designing feedback to induce safer driving behaviors: A literature review and a model of driver-feedback interaction. *Technical Report Submitted to Toyota Collaborative Safety Research Center (CSRC)*.
Gianfranchi, E., Sarlo, M., & Tagliabue, M. (2017). Rischiare in sicurezza: La conduttanza cutanea come marcatore somatico del rischio alla guida in ambiente virtuale. In *In Congresso Nazionale AIP della Sezione Di Psicologia Sperimentale* (p. 134).
Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
Kahneman, D., & Frederick, S. (2005). *A model of heuristic judgment. The Cambridge Handbook of Thinking and Reasoning*. New York: Cambridge University Press.
Kinnear, N., Kelly, S. W., Stradling, S., & Thomson, J. (2013). Understanding how drivers learn to anticipate risk on the road: A laboratory experiment of affective anticipation of road hazards. *Accident Analysis & Prevention, 50*, 1025–1033.
Kluger, A. N., & De Nisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.
Maldonado, A., Serra, S., Catena, A., Cándido, A., & Megías, A. (2016). Modifying evaluations and decisions in risky situations. *The Spanish Journal of Psychology, 19*, E53.
Maldonado, A., Torres, M. A., Catena, A., Cándido, A., & Megías-Robles, A. (2020). From riskier to safer driving decisions: The role of feedback and the experiential automatic processing system. *Transportation Research Part F: Traffic Psychology and Behaviour, 73*, 307–317.
Megías, A., Cándido, A., Maldonado, A., & Catena, A. (2018). Neural correlates of risk perception as a function of risk level: An approach to the study of risk through a daily life task. *Neuropsychologia, 119*, 464–473.
Megías, A., Cortés, A., Maldonado, A., & Cándido, A. (2017). Using negative emotional feedback to modify risky behavior of young moped riders. *Traffic Injury Prevention, 18*(4), 351–356.
Megías, A., Maldonado, A., Cándido, A., & Catena, A. (2011). Emotional modulation of urgent and evaluative behaviors in risky driving scenarios. *Acciden tAnalysis & Prevention, 43*(3), 813–817.
Megías, A., Navas, J. F., Petrova, D., Cándido, A., Maldonado, A., García-Retamero, R., & Catena, A. (2015). Neural mechanisms underlying urgent and evaluative behaviors: An fMRI study on the interaction of automatic and controlled processes. *Human Brain Mapping, 36*(8), 2853–2864.
Megías, A., Torres, M. A., Catena, A., Cándido, A., & Maldonado, A. (2018). Electrophysiological brain indices of risk behavior modification induced by contingent feedback. *International Journal of Psychophysiology, 124*, 43–53.
Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review, 11*(6), 988–1010.
Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13*(2), 60–66.
Sanfey, A. G., & Chang, L. J. (2008). Multiple systems in decision making. *Annals of the New York Academy of Sciences, 1128*, 53–62.
Scialfa, C. T., Borkenhagen, D., Lyon, J., Deschênes, M., Horswill, M., & Wetton, M. (2012). The effects of driving experience on responses to a static hazard perception test. *Accident Analysis & Prevention, 45*, 547–553.
Scialfa, C. T., Borkenhagen, D., Lyon, J., & Deschênes, M. (2013). A comparison of static and dynamic hazard perception tests. *Accident Analysis & Prevention, 51*, 268–273.
Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis: An International Journal, 24*(2), 311–322.
Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research, 177*(3), 1333–1352.
Tagliabue, M., Gianfranchi, E., & Sarlo, M. (2017). A first step toward the understanding of implicit learning of hazard anticipation in inexperienced road users through a moped-riding simulator. *Frontiers in Psychology, 8*, 768.

Tagliabue, M., & Sarlo, M. (2015). Affective components in training to ride safely using a moped simulator. *Transportation Research Part F: Traffic Psychology and Behaviour, 35*, 132–138.
Tagliabue, M., Sarlo, M., & Gianfranchi, E. (2019). How can on-road hazard perception and anticipation be improved? Evidence from the body. *Frontiers in Psychology, 10*, 167.
Tagliabue, M., Gianfranchi, E., & Sarlo, M. (2017). A first step toward the understanding of implicit learning of hazard anticipation in inexperienced road users through a moped-riding simulator. *Frontiers in Psychology, 8*, Article 768.
Torres, M. A., Megías, A., Catena, A., Cándido, A., & Maldonado, A. (2017). Opposite effects of feedback contingency on the process of risky decisions-making. *Transportation Research Part F: Traffic Psychology and Behaviour, 45*, 147–156.
Tūskė, V., Šeibokaitė, L., Endriulaitienė, A., & Lehtonen, E. (2019). Hazard perception test development for Lithuanian drivers. *IATSS Research, 43*(2), 108–113.
Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety Science, 41*(5), 427–443.
Wetton, M. A., Horswill, M. S., Hatherly, C., Wood, J. M., Pachana, N. A., & Anstey, K. J. (2010). The development and validation of two complementary measures of drivers' hazard perception ability. *Accident Analysis & Prevention, 42*(4), 1232–1239.
World Medical Association – WMA. (2008). *Declaration of Helsinki*. World Medical Association.