The Development of Computerised Success-Failure Manipulation Paradigms for the Experimental Study of Metacognition in Neurological Patients

Daniel C. Mograbi

Departamento de Psicologia da Pontifícia Universidade Católica do Rio de Janeiro,
Rio de Janeiro, Brasil

Department of Psychology, King’s College London, London, England

Richard G. Brown
Andrew Brand
Robin G. Morris

Department of Psychology, King’s College London, London, England

Abstract

Research into metacognitive abilities of clinical populations has indicated pervasive deficits in terms of monitoring of cognitive performance. Nevertheless, there are important methodological issues regarding the validity of these findings. In the current paper, we describe the development of novel experimental procedures to mitigate some of these limitations. Specifically, we report the creation of computerised tasks based on success-failure manipulation (SFM), which allow experimental control over the performance of participants. We discuss the theoretical implications of this new procedure and also results obtained so far with the tasks. Finally, we present future research directions stemming from the use of the tasks.

Keywords: Metacognition, awareness, success-failure manipulation, dementia.

Desenvolvimento de Paradigmas Computadorizados de Manipulação de Sucesso-Fracasso para o Estudo Experimental da Metacognição em Pacientes Neurológicos

Resumo

Pesquisas sobre habilidades metacognitivas em grupos clínicos indicaram uma série de déficits em termos de monitoramento de desempenho cognitivo. No entanto, existem importantes questões metodológicas que questionam a validade destes achados. Neste artigo, descrevemos o desenvolvimento de procedimentos experimentais novos para minorar algumas destas limitações. Especificamente, reportamos a criação de tarefas computadorizadas baseadas in manipulação de sucesso-fracasso, que permitem o controle experimental sobre o desempenho de participantes. Discutimos as implicações teóricas deste novo procedimento, bem como os resultados obtidos até então com estas tarefas. Finalmente, indicamos futuras direções de pesquisa partindo do uso destas tarefas.

Palavras-chave: Metacognição, consciência, manipulação de sucesso-fracasso, demência.

1 Mailing address: Institute of Psychiatry, Department of Psychology, King’s College London, PO Box 078, De Crespigny Park, SE5 8AF, London, UK. E-mail: daniel.mograbi@kcl.ac.uk, richard.g.brown@kcl.ac.uk, andrew.brand@kcl.ac.uk and robin.morris@kcl.ac.uk

The success-failure tasks are available free of charge for research purposes, in English or Brazilian Portuguese versions. Requests for the use of the tasks should be made to the corresponding author.
Desarrollo de Paradigmas Computadorizados de Manipulación de Éxito-Fracaso para el Estudio Experimental de la Meta-Cognición en Pacientes Neurológicos

Resumen

Investigaciones sobre habilidades meta-cognitivas en grupos clínicos han indicado una serie de déficits en términos de monitoreo (o monitorización) de desempeño cognitivo. Sin embargo hay cuestiones metodológicas importantes que cuestionan la validez de estas descubiertas. En este artículo, describimos el desarrollo de procedimientos experimentales nuevos para aminorar algunas de esas limitaciones. Específicamente, reportamos la creación de tareas computadorizadas basadas en la manipulación de éxito-fracaso, que permiten el control experimental sobre el desempeño de participantes. Discutimos las implicaciones teóricas de este nuevo procedimiento, así como los resultados obtenidos hasta ahora con estas tareas. Finalmente, indicamos futuras direcciones de investigación partiendo de el uso de estas tareas.

Palabras clave: Meta-cognición, conciencia, manipulación de éxito-fracaso, demencia.

Metacognition can be defined as the knowledge and cognition about cognitive phenomena, including also beliefs and attitudes regarding cognition (Flavell, 1979). In addition, metacognition also refers to monitoring of cognitive performance and use of strategies to achieve goals during cognitive tasks. Due to its potential clinical and theoretical relevance (Morris & Hannesdottir, 2004), a number of different approaches to explore metacognition – also referred in this context as awareness of cognitive abilities – have been employed with psychiatric and neurological patients. Results tend to indicate pervasive deficits in terms of monitoring of cognitive performance in conditions such as dementia (Souchay, 2007), psychosis (David, Bedford, Wiffen, & Gilleen, 2012) and hemiplegia (Orfei et al., 2007). Nevertheless, there are important issues regarding the validity of these findings.

In particular, metacognitive paradigms were originally developed for the study of healthy participants (e.g. school children; Flavell, 1979), and their use with clinical populations brings a number of methodological challenges which are seldom addressed by researchers. Most metacognition measures, such as feeling-of-knowing (FOK), judgement-of-learning (JOL) and judgement-of-confidence (JOC), are based on participant appraisal of performance before/after cognitive tasks. In the case of comparisons between patient groups and controls, three main problems result from this approach. Firstly, heterogeneity in clinical presentation leads to varying levels of performance within patient groups, with failure in tasks ranging from sporadic to constant depending on factors such as disease severity, pattern of brain damage and presence of comorbidities. While some degree of performance variation is also inevitable in healthy samples, it tends to be much higher in conditions such as Alzheimer’s disease (AD), in which the clinical presentation can be quite distinct even in cases at a similar stage of disease severity (Morris & Becker, 2004). This within-group variation in terms of performance may impact estimations of cognitive ability. Secondly, imbalances between patient and healthy groups in terms of actual performance on tasks may drive differences in metacognitive ability. It has been suggested that overestimation of cognitive ability in brain-injured patients may be a statistical consequence, an artefact, of lower scores in performance relative to controls (Marcel, Tegner, & Nimmo-Smith, 2004). Thirdly, because of the presence of cognitive impairments, patients are rarely in a position in which they have high levels of performance during tasks exploring metacognition. This leads to an obvious, but
often overlooked, bias in the literature, namely that little is known about metacognitive ability of successful performance in clinical groups. Investigating metacognitive ability under success conditions may provide important insights about this capacity which are hitherto left unexplored (Mograbi, Brown, Salas, & Morris, 2012).

In the current paper we describe the development of novel experimental procedures to mitigate some of these limitations. Specifically, we report the creation of computerised tasks based on success-failure manipulation (SFM), which allow experimental control over the performance of participants. We discuss the theoretical implications of this new procedure and also results obtained so far with the tasks. Finally, we present future research directions stemming from the use of the tasks.

Overview and Historical Origins of Success-Failure Manipulation

In a typical SFM paradigm, participants are exposed to tasks for which either the level of difficulty of task or feedback about performance is controlled (Nummenmaa & Niemi, 2004). This is frequently in the context of attempting to induce controlled failure for the purposes of either negative mood induction or investigating attributional style in response to failure. Hence, SFM tasks have included, for example, “bogus” intelligence tests, tests of cognitive abilities and social perception skills (Nummenmaa & Niemi, 2004). In terms of mood induction, SFM has the strength that the person participates in the affect eliciting situation, in comparison to other mood induction procedures (MIPS), such as viewing emotional films, or listening to stories in which the participants have a passive role. Another strength of SFM is the fact that, contrary to some other types of MIP (Westermann, Spies, Stahl, & Hesse, 1996), it is equally effective to induce both valences of mood (positive and negative; Nummenmaa & Niemi, 2004). As mentioned above, there are two main methods of SFM: controlling feedback or performance levels. Nummenmaa and Niemi (2004) warn about the problems of feedback manipulation; false feedback may make participants suspicious of the true meaning of the task, hence manipulation of task difficulty should be preferred over giving sham feedback. In addition, Elkin, Whelan, Meyers, Phipps and Glaser (1998) indicate that feedback may not generate robust mood induction effects.

SFM were initially developed to investigate self-esteem in the 1970s and 80s, exploring experimentally the effects of failure on mood and motivation. For example, SFM tasks were used extensively to investigate learned helplessness in humans, as an analogue of experimental procedures used in animals. The term learned helplessness was first used to describe the impaired performance of animals in a training situation produced by prior exposure to uncontrollable aversive stimulation, such as electric shocks (Overmier & Seligman, 1967; Seligman & Maier, 1967). This paradigm was adapted for use in humans (for a review, Roth, 1980), for example, employing impossible tasks, and its theoretical implications were explored to explain mood disorders, such as depression (Abramson, Seligman, & Teasdale, 1978).

In addition to confirming experimentally that failure leads to negative mood, while success results in mild positive/neutral mood (for a review, see Goodwin & Williams, 1982), research about this theme has indicated that effects of failure are stronger for emotions implicating the self (Brown & Dutton, 1995). Some studies suggested that attributional style (e.g. whether the cause of the performance is perceived to be internal or external to the subject) is an important factor in emotional responses to failure, but there is evidence suggesting that valence of the outcome (i.e. success or failure) is the best predictor of affective responses, more than attributions (Whitley, 1986). Research has indicated that some affects are linked with success and failure regardless of the attribution for the outcome. Examples of these outcome-dependent/attribution-independent affects are pleasure, happiness and satisfaction for success, and sadness, displeasure and disappointment for failure (Weiner, Russell, & Lerman, 1979; but see
McFarland & Ross, 1982). It is important to note that the relationship between mood and reaction to performance is bi-directional, with depressive mood also influencing reaction to success – for example with less response to success feedback in depressed patients (Ingram, Smith, & Brehm, 1983). Regarding effects on motivation, there is evidence that experience of failure, in particular repeated failure, leads to less persistence, with less time spent repeating a task after failure, although this effect may also be influenced by self-esteem (Baumeister & Tice, 1985; Shrauger & Sorman, 1977).

**A Novel Use for Success-Failure Manipulation**

Despite their initial use as MIPs, SFM have a number of characteristics which make this methodology a powerful tool for investigations about metacognition and awareness of performance. Firstly, under SFM paradigms performance can be manipulated, bringing the phenomenon of failure under experimental control. Specifically, SFM is about manipulating the level of success or failure in a group of participants, usually to explore the effects of high levels of failure, induced by rigging the task. A modification of this approach, which we have developed, is to systematically modify levels of difficulty so as to match failure or success rates between participants, which we term here as a titration process. By keeping the level of success between participants similar, regardless of actual ability, direct comparisons between clinical and healthy groups become possible, avoiding the confounding effects of different levels of performance (Marcel et al., 2004). In addition, SFM then provides an occasion to study awareness of success performance, which can be infrequent in clinical groups under classical experimental testing conditions. Understanding how metacognition works in successful situations provides important insights about malfunctioning of metacognitive appraisal.

Secondly, this methodology provides an opportunity to measure metacognitive ability, while at the same time engaging participants in an emotionally salient task, such that the effects of failure experience can be explored in relation to awareness of performance. Although it is possible to explore naturally occurring mood in response to failure, this is impractical to study and impossible to manipulate in an experimental situation. In addition, in a SFM paradigm the true nature of the task is disguised, and it is possible to explore covert affective change, while also reducing the risk of demand characteristics that is associated with research into emotion.

Thirdly, SFM tasks have good ecological validity. Given that there is a deception element in SFM tasks, affective responses generated to the experience of failure in them is no different from those caused by failure in regular cognitive tasks. Moreover, SFM resembles real-life situations, since feelings of success and failure are constantly experienced in everyday life, particularly in the case of brain injury or illness. In addition, this sort of procedure allows direct testing of awareness associated with performance on tasks, providing an opportunity for investigating predictions of models of awareness in neurological patients (Mograbi, Brown, & Morris, 2009; Morris & Mograbi, 2013).

**Development of the Tasks**

To control levels of task performance while also inducing mood states, a series of paradigms were developed that systematically exposed participants to either success or failure. Since no previous paradigms of SFM have been used before with neurological patients, novel procedures were developed, also incorporating a titration method. After piloting, further modifications were made leading to the final procedure, as described below.

The current tasks are presented via a computer program, the program providing easy and systematic adjustment of difficulty levels according to participant performance. Tasks were programmed using REALbasic 2007 (Release 5). For each task the participant’s personal performance threshold (i.e. the point when he/she
started to fail on the task) is determined, and then, according to the experimental condition (success or failure), trials are presented above or below this threshold. The level of success or failure and the amount of time on each phase can be set by the experimenter, as well as the order of procedures. Participants are not aware of the fact that the tasks have a set level of difficulty. All tasks have a preparatory phase to familiarise participants with the procedures and a testing phase, which is further divided into a titration and experimental phase. We have developed a total of four tasks divided into two types, according to cognitive domain: reaction time and memory tasks. For each task type, there are two parallel versions developed, allowing the use of different experimental conditions (e.g. success or failure) for each domain while also controlling for specific task-related effects. However, tasks need to be visually distinct enough so the effects of each condition are discernible.

**Reaction Time Tasks**

In the reaction time tasks, participants are given a button box with a single centrally located button. In the first task, in each trial they have to press the button as soon as they see an animated moving car appear on the screen. In the first version, the screen shows a road and any time after a sound warning, a car (drawing) appears moving from left to right (for screenshots in a reduced frame, see Figure 1). The button has to be pressed before the car reaches the right hand side of the screen, the speed of the car used to manipulate task difficulty level. Car models and colours vary in each trial. In the second version, participants have to press the button as soon as they see an object appearing on the screen falling from top to bottom. The screen shows a vertical grey bar at the left (looking like the side of a building) and any time after a sound warning an object (pictures; e.g. a ball, an egg, a vase) appears to the right of the bar, falling from the top to the bottom (for screenshots in a reduced frame, see Figure 2). Here the participant has to ‘catch’ the object before it gets to the bottom by pressing the button, again with the falling speed used to determine difficulty. Objects are different in each trial. In both versions, feedback is given; if participants press the button on time, respectively, a traffic warden (drawing) appears stopping the car or a hand (picture) appears catching the objects, and a ‘clink’ is heard; otherwise the car or object just keeps moving and a croaky ‘beep’ is heard indicating failure.

**Memory Tasks**

The first version is a visual span task. In each trial, participants have to pay attention to the screen, which displayed a number (from 1 to 10) of identical everyday objects (e.g. alarm clocks, baskets, teapots). The objects are scattered in random position on a light blue “plinth” (for screenshots in a reduced frame, see Figure 3). A red square then appears around one of the objects highlighting it and randomly moves to other objects on the set. The minimum sequence highlights a single object and the maximum 10 objects, the number for each trial used to determine difficulty level. After the sequence is over, the objects remain on the same position and participants have to point the sequence in which the objects have been highlighted. The experimenter then clicks on the objects in the sequence pointed by the participant using a mouse (touch sensitive screen technology may also be used). The second version is a digit span task, in which participants have to listen to a sequence of digits and immediately repeat it back to the experimenter. Digits, ranging from 0 to 9, appear in black, at the centre of the white screen (for screenshots in a reduced frame, see Figure 4), and are read out loud by a native English speaker (a Brazilian Portuguese version has also been developed). The shortest list-length is a single digit and the longest ten digits, with difficulty being controlled by adjusting list-length. Feedback about performance is automatically given in the form of a green tick and a ‘clink’ for a correct answer or a red cross and croaky ‘beep’ for a mistake.
All tasks have the same following phases:

**Practice**

There are a configurable number of trials with low-difficulty levels for each task, allowing participants to familiarise themselves with the tasks and giving the opportunity for the experimenter to give further instructions.

**Titration Phase**

After practising, the actual task starts. The first trials of the task serve to establish the par-
participant’s individual reaction time or memory threshold. This essentially is how well the participant is capable of doing that specific test, with the notion that if the task is made harder than this level they are likely to fail consistently and if it is made easier they are likely to succeed. Participants are not informed of the purpose of this phase and there is no apparent discontinuity between that stage and the following phases. The trials increase in difficulty until the participant fails consecutive trials (configurable number) at the same difficulty level, this level defining their individual threshold.

**Experimental Phase (Success/Failure Manipulation)**

In this phase, participants are presented with trials above or below their threshold, according to the testing condition (success or failure). Control of performance is done with blocks of trials; for each 10 trials, the experimenter configures for a particular experiment how many will be above (failure condition) or below (success condition) the participant’s threshold, giving full control over difficulty levels. Duration of this phase is configurable, and the program includes an option of having two different experimental phases in the same task, with no apparent discontinuity between them, allowing the comparison of different success/failure conditions within the same version. Alternatively, the two versions for each type of task can be used to compare success and failure conditions.

**Pilot Work**

Development of the computer tasks went through two pilot stages. In the first stage, the program was piloted in a small group of young adults \((n = 6)\). This first testing was important for technical adjustments, such as the control of the difficulty level of the tasks. For example, it was found that when participants started the success or failure phase, their ability level might change and this compromised the originally titrated thresholding for each participant. Accordingly, a monitoring element was introduced, such that if participants’ performance improved or worsened, the program would adjust the level of difficulty to keep the expected success/failure constant. This was facilitated by there being computer control for both types of tasks. Another important change was the introduction of features to pause and restart trials, to avoid interruptions/distractions affecting performance and also to allow further instructions to participants during the tasks if necessary. In the second stage of piloting, the revised version of the software was tested in a small group of participants, including neurological patients (6 older adults and 4 AD patients). This testing was important to determine if participants were able to fully understand the computer tasks. Pilot testing indicated that participants were able to attend appropriately to instructions and do the tasks, not showing significant comprehension problems about the procedures, despite the presence of cognitive deterioration in the case of AD patients (for published results with the tasks, see Mograbi et al., 2012). Finally, both pilot stages helped to establish that mood states generated by the tasks were mild and short lived (an important ethical requirement), to determine the optimal length of break between tasks and to confirm that the true purpose of the tasks was not identifiable.

**Efficacy of the Tasks**

The SFM procedure has now been successfully used in two studies exploring awareness and emotional reactivity/behavioural adaptation in AD. Mograbi et al. (2012) investigated experimentally induced emotional responsiveness in relation to task failure in an AD group \((n=23)\) and matched controls \((n=22)\). The SFM paradigm enabled matching error rates between groups, necessary to explore emotional reaction given the same degree of failure. Half of the tasks were rigged to be above participants’ ability level (success tasks), while the other half was below (failure tasks). Results indicated that, although AD groups had less awareness of failure relative to controls, emotional reactivity was preserved, both in terms of self-report (emotions such as frustration, disappointment and embarrassment) and filmed facial expressions, with
increased reactivity to failure compared with success. In all tasks, emotional reactivity to failure was not correlated with awareness of performance, which the authors interpret in terms of implicit awareness. Importantly, this was the first study to show that awareness of success performance is also impaired in AD, suggesting a central problem in terms of monitoring which impairs performance evaluation as whole, without a positive bias.

A second study (Mograbi et al., submitted) explored long-term reactivity in a follow-up session with the AD participants from the study described above. The participants involved in the first study were followed up a week later, and this time, all tasks (i.e. whether success or failure) had a 50% success level. Tasks had no time limit, and participants had to indicate when they wanted to stop doing them, with the hypothesis being that success or failure in the first session would lead to relative changes in persistence. Participants were more likely to give up on previous tasks in which there was systematically induced failure, suggesting an implicit learning process. This was in the context of the AD patients having little or no memory doing the tasks previously. Interestingly, participants who tended to persist in doing (previously) success tasks tended to show better awareness of performance as measured in the first session.

Altogether, these two studies demonstrate the efficacy of the SFM paradigm in controlling performance, while also illustrating its flexibility to explore topics related to metacognition, awareness, emotional reactivity and adjustment to performance. It is important to note that participants, also including controls, did not express any suspicion that performance levels were being rigged to explore responses to success or failure. Also, the task was reviewed ethically with respect to being used for potentially more vulnerable adults; it was considered that mood alterations generated by exposure to failure with these tasks were transient and within the range of ordinary daily experience. Additionally, features of the study were designed to reassure patients and ensure an overall positive experience (e.g. positive mood induction and debriefing at the nature of the tasks at the end of the testing session).

Future Research

A main future research direction is to explore the potential of the tasks to bring levels of performance under experimental control to investigate metacognition and awareness in healthy and clinical populations. For example, the SFM tasks can be applied to other clinical populations, now that it has been shown to be a useful methodology to explore metacognition in people with early-stage dementia. For instance, this could be carried out in other conditions in which awareness of performance may be compromised, such as in patients with varying degrees of hemiparesis/hemiplegia severity (Orfei et al., 2007); because the reaction time tasks are motor in nature, and the level of difficulty on them can be configured, these may be useful tools for understanding the presence of impairments of awareness of motor performance in this condition. Although the tasks may not be suitable for all patients with this condition, either because of plegia/paresis or anosognosia severity, the methodology is flexible enough to be adapted for this context. In addition, it may be especially interesting to use the computer tasks with patients with frontotemporal dementia, who are known to have problems with error monitoring (Snowden, Neary, & Mahn, 2002).

Systematic control of performance in SFM paradigms also make them suitable for the exploration of the neural correlates of metacognition. For example, the tasks may be particularly suited for functional neuroimaging studies which typically involve blocks of contrasting events (e.g. success and failure phase of the tasks), with a discrete beginning and end (such as the trials of the tasks). Likewise, electrophysiology studies about error monitoring can also benefit from the systematic and structured presentation of failure trials in the developed tasks.

Finally, the tasks can also be used as a standalone mood induction procedure, and they proved robust in this respect. Meta-analyses have indicated the effectiveness of SFM as a
mood induction procedure, and its ecological validity in relation to other methods has also been suggested (Nummenmaa & Niemi, 2004). In particular, the tasks may be useful to induce affective responses related to cognitive failure, such as frustration, disappointment and embarrassment (and, conversely, positive emotions such satisfaction and pride).

Concluding Remarks

To conclude, we believe that the tasks developed can help to mitigate a series of limitations currently associated with studies on metacognitive ability of clinical populations, while also opening up new research possibilities not foreseen before, as outlined above. More generally, we believe that neuropsychology has much to gain when well established research paradigms from the past can be revitalised by technological approaches, expanding their scope and applicability.

References


Xojo Inc. (2007). REALbasic (Release 5) [Computer software]. Austin, TX: Xojo Inc.


Received: April, 15, 2013
1st revision: October, 03, 2013
Accepted: October, 16, 2013