Effectiveness and feasibility of Socratic feedback to increase awareness of deficits in patients with acquired brain injury

NEUROPSYCHOLOGICAL REHABILITATION
2020, VOL. 30, NO. 4, 591–612
https://doi.org/10.1080/09602011.2018.1485110

Effectiveness and feasibility of Socratic feedback to increase awareness of deficits in patients with acquired brain injury: Four single-case experimental design (SCED) studies

Anne-Claire M. C. Schrijnemaekers[a], Ieke Winkens[b], Sascha M. C. Rasquin[a,c], Annette Verhaeg[a], Rudolf W. H. M. Ponds[a,d,e] and Caroline M. van Heugten[b,d]

[a]Department of Brain Injury Rehabilitation, Adelante Rehabilitation Centre of Expertise in Rehabilitation and Audiology, Hoensbroek, Netherlands; [b]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands; [c]Department of Rehabilitation Medicine, CAPHRI School for Public Health and Primary Care, Maastricht University, Hoensbroek, Netherlands; [d]Department of Psychiatry and Neuropsychology, Faculty of Health, Medicine and Life Sciences, School for Mental Health and Neuroscience, Maastricht University, Hoensbroek, Netherlands; [e]Department of Medical Psychology, Maastricht University Medical Centre, Hoensbroek, Netherlands

ABSTRACT

Objective: To investigate the effectiveness and feasibility of a Socratic feedback programme to improve awareness of deficits in patients with acquired brain injury (ABI). Setting: Rehabilitation centre. Participants: Four patients with ABI with awareness problems. Design: A series of single-case experimental design studies with random intervention starting points (A–B + maintenance design). Main measures: Rate of trainer-feedback and self-control behaviour on everyday tasks, patient competency rating scale (PCRS), self-regulating skills interview (SRSI), hospital anxiety and depression scale. Results: All patients needed less trainer feedback, the change was significant in 3 out of 4. One patient increased in overt self-corrective behaviour. SRSI performance increased in all patients (medium to strong effect size), and PCRS performance increased in two patients (medium and strong effect size). Mood and anxiety levels were elevated in one patient at the beginning of the training and decreased to normal levels at the end of the training. The feasibility of the programme was scored 9 out of 10. Conclusions: The Socratic feedback method is a promising intervention for improving awareness of deficits in patients with ABI. Controlled studies with larger populations are needed to draw more solid conclusions about the effect of this method.

ARTICLE HISTORY Received 2 May 2016; Accepted 29 May 2018

KEYWORDS Awareness of deficits; Acquired brain injury; Socratic feedback training; Single-case experimental design

Introduction

Impaired awareness of deficits is common among persons who have experienced an acquired brain injury (ABI) and can be present for more than 5 years post-ABI (Kelley

CONTACT Anne-Claire M.C. Schrijnemaekers a.schrijnemaekers@adelante-zorggroep.nl

Supplemental data for this article is available from doi:10.1080/09602011.2018.1485110.

et al., 2014). The reported incidence within this population ranges from 30% to 97% (Bach & David, 2006; Flashman & McAllister, 2002; Goverover, Johnston, Toglia, & DeLuca, 2007) and depends on multiple factors, including the measurement instrument used (Fleming, Strong, & Ashton, 1996), the severity of the injury (Bach & David, 2006; Vanderploeg, Belanger, Duchnick, & Curtiss, 2007), and the time since the injury (Vanderploeg et al., 2007). Areas of functioning in which patients may lack awareness include cognitive, behavioural, emotional, and interpersonal skills as well as motor and sensory functioning (Bogod, Mateer, & Macdonald, 2003; Hart, Seignourel, & Sherer, 2009; Port, Willmott, & Charlton, 2002).

Crosson et al. (1989) defined impaired awareness of deficits in a three-stage hierarchical model in which intellectual awareness represents the ability to understand that a function is impaired or to understand the complications caused by the impairment. Emergent awareness is the ability to recognise a problem when it is occurring. Anticipatory awareness is the ability to foresee that a problem will occur as a result of having a deficit. Other models concerning impaired awareness of deficits after ABI show considerable overlap with one another (Allen & Ruff, 1990; Barco, Crosson, Bolesta, Werts, & Stout, 1991; Fleming et al., 1996; Giacino & Cicerone, 1998; Langer & Padrone, 1992; Sohlberg & Mateer, 2001). Often, these models include cognitive components (basic cognitive abilities to perceive, understand and recognise situations), metacognitive components (self-regulation and anticipation), and psychological components (with denial on one side of the spectrum and acceptance on the other) (Schrijnemaekers, Smeets, Ponds, van Heugten, & Rasquin, 2014).

Self-regulation is a central aspect because being able to function independently across different activities and contexts requires that one is able to monitor their performance and knows when a particular strategy should be applied (Toglia, Johnston, Goverover, & Dain, 2010). Knowing when a particular strategy should be applied depends on whether similarities between new and previous experiences are recognised. Only then are learned abilities activated (Bransford, Brown, Cocking, Donovan, & Pellegrino, 2000). Self-regulation can be measured with interviews, self-reports and direct measurement procedures (Ownsworth, Quinn, Fleming, Kendall, & Shum, 2010), such as observing error detecting and correction (Hart, Giovannetti, Montgomery, & Schwartz, 1998).

Impaired awareness of deficits may diminish motivation for treatment, as the patient may not recognise or acknowledge the need for intervention (Flashman & McAllister, 2002). Hence, lower treatment adherence and performance can be expected, which in turn can lead to unfavourable short- and long-term employment outcomes (Sherer et al., 1998; Trudel, Tryon, & Purdum, 1998), behavioural problems (Trudel et al., 1998) and higher caregiver burden (Kelley et al., 2014). The adequate treatment of impaired awareness is therefore recommended.

Currently, no standard treatment protocol exists to improve awareness in patients with ABI. The small sample sizes in studies as well as the heterogeneity of the interventions makes conclusions only tentative (Schmidt, Lannin, Fleming, & Ownsworth, 2011; Schrijnemaekers et al., 2014). Interventions that consist of multiple components or those that are not specifically described make it difficult to disentangle the training elements that were successful in improving awareness. Schmidt et al. (2011) concluded, based on three randomised controlled trials (RCT), that it is not possible to determine from these studies whether one form of feedback is more effective than another. Schrijnemaekers et al. (2014) concluded, based on nine high quality studies (2 RCT, 5 SCED, 1 single-case design with follow-up, 1 quasi experimental controlled design), that two combined components appeared to be effective in improving awareness in ABI patients in the post-acute and chronic phase: (1) teaching functional skills in multiple settings with (2) feedback using guided experience in a non-confrontational manner, such as the Socratic dialogue.

In the present study, we used a series of single-case experimental design (SCED) studies to evaluate the feasibility and effectiveness of an awareness intervention programme based on the Socratic feedback method for multiple daily cognitive tasks. Socratic questioning is a demonstrated technique from cognitive behavioural therapy (CBT) for the treatment of depression and anxiety (Roth & Pilling, 2007). Therapists use questions to foster active engagement and critical thinking, thereby aiding in the learning process (Neenan, 2009) and helping patients develop new perspectives (Overholser, 1993; Padesky, 1993). The aims of the present study were to investigate the effectiveness and feasibility of Socratic feedback to increase awareness in terms of improvements in specific self-regulatory behaviour, and general levels of self-regulation, awareness of deficits, depression and anxiety.

Methods

Participants

Four ABI patients (three patients post-stroke, one patient with traumatic brain injury) with an impaired awareness of cognitive deficits participated. A lack of awareness was based on the interpretation of the Prigatano competency rating scale (Fordyce & Roueche, 1986) difference score in combination with the clinical judgement of behaviour by a physician, nurse and psychologist, such as the absence of cognitive complaints in combination with a poor performance on neuropsychological tests.

Inpatients and outpatients in the post-acute phase of ABI (after hospital discharge to 6 months post-injury) and 18 years or older who were treated at Adelante Rehabilitation Centre in Hoensbroek, Netherlands, from April to December 2014, were eligible to participate. Exclusion criteria were post-traumatic amnesia, physically aggressive behaviour and severe problems in the domain of language and perception, all based on clinical judgement.

Design

For each patient, a SCED was used in which the starting point of the intervention was randomised across participants. Patients' symptoms were evaluated over time across baseline (A), intervention (B) and maintenance (m) phases. Baseline symptoms were evaluated before the start of the intervention. The start of the treatment phase was determined randomly per participant, using the website www.randomizer.org (Urbaniak & Plous, 2013), given the restriction that the baseline phase (A) should last for at least 2 weeks (8 sessions) and at most, 7 weeks (28 sessions). The minimum length of the baseline was set to provide an adequate baseline measurement (Tate et al., 2013); the maximum length of baseline was based on practical reasons. The intervention phase (B) consisted of a fixed number of 16 sessions (4 weeks) of treatment. The maintenance phase (m) also included 16 sessions for all four patients. In the maintenance phase the behaviour is expected and desired to be similar to the B phase level.

To improve internal validity and statistical conclusion validity, the starting point of the intervention and therefore the length of the baseline period was randomised (Edgington & Onghena, 2007). Randomisation strengthens the internal validity of SCEDs because it yields statistical control over known and unknown confounding variables (Heyvaert & Onghena, 2013; Onghena, 1992). For statistical reasons alone, multiple assessments were performed during the baseline, intervention and maintenance phases. A pre-intervention (baseline) performance could therefore be established and patterns of performance could be compared between different phases (Kratochwill & Levin, 2010).

Training

The intervention took place at the Adelante Rehabilitation Centre in a neutral and quiet therapy room 3×4 m^2, with a window, table, some chairs, paper and pencil equipment, and computer facilities. Each session lasted 60 minutes. Functional and compensatory training of cognitive deficits occurred in all phases (A, B, m). The cognitive trainer determined the choice and complexity of the training tasks based on the cognitive problems of the patient. Three levels of complexity were distinguished (single simple tasks, single complex tasks and multiple complex tasks), based on the different levels of complexity distinguished by the International Classification of Functioning, described in the "activities and participation" section (WHO, 2001). At the start of the programme the trainer chose the level of complexity, and this level remained unchanged. Within the chosen category, the tasks that fell within the patients' field of interest or were considered by the patient to be important for daily living were chosen. The scores were only used for within-subject comparisons.

In the baseline (A) and maintenance phase (m), the trainer gave feedback as usual (detailed information can be found online in Appendix 1). This feedback could best be described as a combination of different feedback styles, as is often used in the clinic. It included correcting inadequate approaches, providing positive verbal reinforcement for accurate responses and explicitly pinpointing inadequate responses. In the intervention phase (B), Socratic feedback was given. The Socratic method is a cognitive restructuring technique grounded in Ellis's rational emotive therapy (Ellis, 1962) and Beck's cognitive therapy (Beck, 1967). The Socratic method is defined as a dialogue between therapist and patient in which the former tries to make the patient reflect on the appropriateness of their cognitions. This is followed by a process of shaping, through which the therapist modifies the patient's verbal behaviour in the overall direction of their chosen therapeutic objectives by repeatedly asking questions (Calero-Elvira, Froján-Parga, Ruiz-Sancho, & Alpañés-Freitag, 2013). The Socratic feedback dialogue differed from the feedback during the baseline in that the trainer repeatedly and consistently used the Socratic feedback method, instead of multiple forms.

The Socratic feedback method that we derived from the Socratic method consisted of asking one or more of the following types of open questions that urge the patient to reflect on their performance: How do you do it? Why does it go like this? What do you think of it? What does it mean for you and your goals? (Detailed information can be found online in Appendix 1). When the patient becomes stuck (because s/he does not have the answer or solution, or s/he makes an error) the cognitive trainer pauses and waits for the patient to self-correct. If that does not happen, the cognitive trainer first encourages the patient, without giving (parts of) the solution, e.g., "What are you looking for?". This stimulates the patient to pay attention and proceed with what s/he was doing. In line with Ownsworth, Fleming, Desbois, Strong, and Kuipers (2006), we call this the non-specific prompt. If the non-specific prompt does not lead to the right answer or solution, the cognitive trainer gives a specific prompt to help the patient find the solution, e.g., "Have you already looked at the right side of the table?". The specific prompt helps the patient in the right direction. If the patient still cannot find the solution or the answer, the cognitive trainer corrects the patient in a neutral way: "Look, this is the right one".

Directly after completing a task, the patient is asked to reflect upon their performance. Then, the trainer gives Socratic feedback about the actual performance (what went right and what went wrong) by encouraging the patient to relate the performance on the task to their personal goals or values and by asking the patient how to reach these goals or live these values. By asking questions and relating this to valuable aspects of the patient's daily living kindly and without negative judgement or persuasion, the patient is encouraged in a safe environment to question some of their own abilities while getting reassurance in other abilities or personal characteristics ("you are a real go-getter", "I like your sense of humour", "I see you are a family-man").

The Socratic feedback method resembles the "Pause, Prompt, and Praise" technique (Glynn, McNaughton, Robinson, & Quinn, 1979; McNaughton, Glynn, & Robinson, 1987), with the difference that with the Socratic feedback method the prompting is mainly through questions. The naming of the non-specific prompts, specific prompts and self-corrections is based on Ownsworth et al.'s (2006) metacognitive contextual intervention on patient JM. Detailed information can be found online in Appendix 1.

Two cognitive trainers participated in this study. Cognitive trainer ML performed all four interventions (phase B) and maintenance training (phase m). Cognitive trainer HM performed the baseline training (phase A) in patients K and S; ML performed the baseline training in patients R and B. Cognitive trainer ML was educated in the Socratic feedback method and trained by the first author in the Socratic method by means of a 90-minute education session. The training of the cognitive trainer included role-play, and frequent evaluations were made during the training of three patients in a pilot study.

Before the four patients described in this study began their training, three other patients participated in a pilot study. The aim of this pilot study was to refine the Socratic feedback method in order to make it applicable for the trainer and to improve the scoring protocol. The feedback method of the cognitive trainer ML was evaluated weekly in order to verify that he used the Socratic feedback method correctly and consistently. As it was not the aim to measure improvements in target measures, the results were not analysed statistically.

Primary outcome: Rate of feedback and self-control behaviour

In this study, we used the rate of trainer-feedback and the rate of self-control behaviour based on, but not copied from, the error behaviour proposed by Ownsworth et al. (2006).

Trainer feedback includes non-specific and specific prompts and corrections. Non-specific prompts refer to general questions or comments that stimulate the patient to pay more attention (e.g., "What were you searching for?", "Do you remember the pitfalls we talked about?"). Specific prompts supply the patient with information about the solution (e.g., "Could you take a look at the left side of the page?", "The name of the movie starts with Ti … "). Corrections refer to the trainer neutrally pointing out a mistake or wrong answer or the trainer giving the right answer (e.g., "No, that is not right", "The name of the movie is Ti … tanic"). The self-control behaviour of the patient includes observable self-corrections and the patient asking the cognitive trainer for help (more information can be found online in Appendix 1).

All training sessions conducted by the cognitive trainers with each participant through the A, B and maintenance phases of the study were videotaped by the cognitive trainer. These videos were used to rate the frequency of trainer non-specific prompts, trainer specific prompts, trainer corrections, client self-corrections and client requests for help.

In the pilot study, we assessed the inter-rater reliability for the scoring system of the trainers' feedback and the self-control behaviour of the patient on multiple tasks. Two observers rated the self-control and feedback behaviour independently on six videotapes of two pilot patients by scoring the rate of self-corrections, requests for help, prompts and corrections during each task, following the scoring protocol. The inter-rater agreement between the two observers for the different types of trainer-feedback and the self-control behaviour fell in the range of good reliability. The intra-class correlation coefficient (ICC) (two-way random model, single measure) was calculated. The inter-rater reliability was excellent (ICC = .89; $p = .00$; $n = 30$). An ICC >.74 is considered excellent, between .74 and .60 good, between .59 and .40 fair, and below .40 is considered poor inter-rater reliability (Cicchetti, 2001).

The rate of corrections, non-specific prompts, specific prompts, requests for help and self-corrections were measured during the performance on multiple everyday cognitive tasks. To differentiate between actual improvement in awareness (less feedback, more self-corrections) and task-specific learning, regular tasks and repetitive tasks were administered. Progress on the regular tasks represents an improvement in awareness, while progress on the repetitive tasks but not on the regular tasks is more likely to represent task-specific learning. This way it becomes clear whether the patient is able to transfer knowledge from one task to other tasks.

In every session, one regular and one repetitive task were performed. The repetitive task was repeated for at least ten consecutive sessions, i.e., in ten consecutive sessions, the exact same task was practised. After these ten sessions, the trainer replaced the old repetitive task with a new repetitive task for at least the next ten sessions. However, the regular task changed every session, i.e., in each session a different but comparable (in terms of content and complexity) task was used. The repetitive tasks were equal to the regular tasks with respect to the choice of subject and cognitive complexity.

Secondary outcome: General awareness, emotional distress

The PCRS developed by Fordyce and Roueche (1986) is a measurement of the general awareness of deficits. The test–retest reliability is acceptable (Fleming et al., 1996; Prigatano, Altman, & O'brien, 1990) and the test has been validated for construct in traumatic brain injury patients (Sherer, Hart, & Nick, 2003). Using a 5-point Likert scale, the patients and the occupational therapist rated on a weekly basis the problems the patient encountered in performing 30 functional activities (Fleming et al., 1996). The PCRS difference score is measured by subtracting the patient score from the therapist score. Scores below 28 indicate no or mild impairments in awareness, and scores between 28 and 51 indicate moderate impairments in awareness. Scores above 51 indicate severe impairments in awareness (Sherer et al., 2003). The self-regulating skills interview (SRSI) is a measure of a range of metacognitive skills and is composed of one screening item and six items that assess different key metacognitive or self-regulation skills regarding a main area of difficulty: intellectual, emergent and anticipatory awareness; readiness to change; strategy generation; degree of strategy use; and strategy (Ownsworth, McFarland, & Young, 2000). For each item, scoring is based on a 10-point Likert scale. For this study, we used the summed scores of the six items, although we acknowledge that this is not the original way of scoring and validity has not been tested for this sumscore. Sumscores could range from 0 to 60 with lower scores corresponding to a higher level of awareness. The inter-rater and test–retest reliability of the SRSI is acceptable, and empirical support regarding the construct and concurrent validity has been reported (Ownsworth et al., 2000). The PCRS and SRSI were assessed weekly, with the day of the week depending on the therapy schedule of the patient.

Improvements in the awareness of deficits are found to be related to increasing levels of emotional distress (Anson & Ponsford, 2006). Hence, mood and anxiety were monitored with the hospital anxiety and depression scale (HADS) (Zigmond & Snaith, 1983). The HADS was administered once every 2 weeks. This 14-item self-report screening scale has good convergent validity in the Dutch population (Pouwer, Van Der Ploeg, Ader, Heine, & Snoek, 1999). Scores between 8 and 10 represent mild symptoms, scores between 11 and 15 represent moderate symptoms and scores of 16 and higher represent severe symptoms (Olssøn, Mykletun, & Dahl, 2005).

Feasibility

Within a week after the last session of the maintenance phase, feasibility questionnaires were completed by each patient and by the cognitive trainer. Questions about the content, duration, intensity, timing, effects and utility (use of strategies learned in everyday life) of the intervention were assessed. The evaluation questionnaires can be found online in Appendix 2.

Procedure

The local ethics review board approved the study. Patients were selected based on convenience sampling. The psychologists of the Department of Brain Injury selected the patients with the most pronounced rehabilitation-interfering impaired awareness of deficits, based on their clinical judgement and the PCRS scores. These patients and their families were informed verbally and in writing about the goal and content of the study by their neuropsychologist. In case of participation, an informed consent was signed. When a start date for the training was available (taking into account trainer availability and holidays), the first patient on the list was informed that the training would begin. All sessions were videotaped by the cognitive trainer and the videotapes were given to the neuropsychologist, who assigned the videotapes to one of four raters, all of whom were masters students in clinical psychology. They were not informed about whether the sessions came from the baseline, intervention or maintenance phase. To ensure that the four raters were consistent in scoring, the students practised together to harmonise counting behaviour in a pilot study with three patients, using a standardised scoring protocol (see the scoring protocol, available online). To monitor consistency of scoring, the accuracy of the rating protocol was evaluated in the first four sessions, the A-phase, the B-phase and the m-phase in the three pilot cases. The protocol was refined during the pilot in patient 1, based on the experience of the raters. In total, 3 × 12 pilot sessions per patient (36 pilot sessions) were evaluated by the neuropsychologist. Feedback was given to the students with respect to rating the scores. After these pilot sessions, the raters felt content with the ease of scoring using the scoring protocol and the scoring was consistent among raters. After these pilot sessions, patients K, B, R and S started the programme.

To apply an accurate procedure, agreements were made between the cognitive trainer and neuropsychologist with respect to the content, application and duration of the tasks, the location of the therapy, and feedback. A formal and numeric evaluation, as described by Ledford and Gast (2014) was not performed. A schedule of tasks that was dependent on the individual cognitive characteristics of the patients and their areas of interest was created in a meeting by the cognitive trainer and a neuropsychologist and included in a report at the start of the baseline phase. The dates and times of the training schedule were organised by an administrator. The accuracy of the protocol was evaluated by the neuropsychologist on a weekly basis. She read the report that the cognitive trainer completed directly after the sessions with the three pilot patients, and looked at one videotape per week in week 1 and two videotapes of all three phases (A–B–m) for all three pilot patients. Twice, the cognitive trainer requested that the neuropsychologist look at another session to receive feedback about that session. If the response of the cognitive trainer was not in line with the Socratic feedback protocol, this was discussed with the cognitive trainer and advice was given and small adjustments to the Socratic feedback protocol were made. Discussions about how to respond according to the protocol occurred most often with respect to pilot patient 1, incidentally in pilot patient 2 and did not occur with respect to pilot patient 3. By then, the cognitive trainer accurately and consistently followed the protocol. No adverse events occurred.

Table 1 shows an overview of the duration, number of sessions and primary and secondary outcome measurements in the baseline, intervention and maintenance phases.

Statistical analysis

The duration of the task performance within the training sessions varied slightly across the different training tasks, as some tasks were shorter than others. To be able to accurately compare the performances in the analyses, the duration of each session was transformed into a "block" of 10 minutes by the formula (rate of target behaviour [prompts, corrections, requests for help, self-corrections] / actual performance time in minutes × 10).

Table 1. Overview of the duration, number of sessions and primary and secondary outcome measurements in the baseline, intervention and maintenance phases.

	Baseline A			Training B				Maintenance m			
Duration	Variable, 2–7 weeks			Fixed, 4 weeks				Fixed, 4 weeks			
Sessions	4	4	. .	4	4	4	4	4	4	4	4
Feedback and self-control rate	4	4	. .	4	4	4	4	4	4	4	4
PCRS, SRSI	1	1	. .	1	1	1	1	1	1	1	1
HADS	1		. .	1		1		1		1	

PCRS = Patient competency rating scale (Fordyce & Roueche, 1986) ; SRSI = Self-regulation skills interview (Ownsworth et al., 2000); HADS = Hospital anxiety and depression scale (Zigmond & Snaith, 1983)

For each patient, the data regarding feedback and control behaviour on the regular and repetitive tasks were plotted graphically to allow for visual inspection. The visualisation of the changes in levels across phases was facilitated by superimposing horizontal reference lines that represented the average scores for the baseline and intervention phases on the raw time series in the plots (Bulté & Onghena, 2012).

A randomisation test (Edgington & Onghena, 2007) was used to assess the statistical significance of the changes in levels across phases per patient. The application of a randomisation test requires some aspect of the experimental design to be randomised. In single-case studies, randomisation cannot be accomplished by randomly assigning subjects to treatments. Instead, measurement occasions can be randomly assigned to treatments. In our case, the point in time at which treatment began was randomly selected before the start of the study for each patient (Edgington, 1980).

With randomisation tests, the smallest possible p-value that can be obtained equals the inverse of the number of possible assignments or permutations. In our study, the use of 21 randomisations permits statistical significance at the $<.05$ level (Bulté & Onghena, 2008). Thus, in order to obtain a p-value of $<.05$, the required minimum number of possible assignments or permutations is 21 ($1/21 = 0.048$). Theoretically, the intervention could start at any observation occasion between 8 and 28 (i.e., not before or at the first 7 observation occasions, and no later than the 29th observation occasion). There were therefore 21 starting points or permutations.

For each patient, we randomly assigned the point in time at which the experimental treatment (Socratic feedback) began. With a randomisation test, whether measurements before the point in time at which experimental treatment began differed significantly from measurements after that point could be evaluated. The null hypothesis is that test scores of the patient are independent of the treatment phase under which they are observed, i.e., no effect of treatment (Edgington & Onghena, 2007). A t-statistic was used to compare mean differences between phase scores. The statistic was used for all the possible permutations of the data generated during the randomisation process: t-statistics were calculated as if data had been collected for each of the 21 permutations. The probability associated with the observed statistic was computed by dividing the number of computed statistics that were as large or larger than the statistic for the observed data set by the total number of possible permutations (Perdices & Tate, 2009).

Effect sizes were calculated for changes in performance on PCRS and SRSI using "nonoverlap of all pairs" (NAP) (Parker & Vannest, 2009). The comparisons performed using the randomisation test and NAP entail the baseline phase (A), the intervention phase (B) and the maintenance phase. In the comparisons the measurements in the maintenance phase were considered as measurements of the B phase, because the behaviour in the maintenance phase is expected and desired to be similar to the B-phase level. NAP summarises data overlap between each phase A datapoint and each phase B datapoint, in turn. A nonoverlapping pair will have a phase B datapoint larger than its paired baseline phase A datapoint. NAP equals the number of comparison pairs showing no overlap, divided by the total number of comparisons. NAP was calculated by hand and interpreted using the following NAP-ranges (Parker & Vannest, 2009): weak effects: .0–.65; medium effects: .66–.92; large or strong effects: .93–1.0.

Results

Four patients were included in this study. The patient characteristics are summarised in Table 2. One patient scored below the PCRS cut-off criterion at baseline. Patient K's answers were, compared to the answers of the therapist, an underestimation in the emotional and interpersonal domain and an overestimation in the cognitive and practical domain. Hence, the many discrepancies in the PCRS answers resulted in a small total discrepancy score that indicated only a mild impairment in awareness.

Primary outcome: Rate of feedback and control behaviour

Randomisation qualified patient S to 16 baseline measurements, and therefore she completed a total of 48 cognitive training sessions (phase means A: 16, B: 16, m: 16). The tasks were related to her hobbies and daily activities such as household chores, cooking, reading and shopping. The total amount of feedback during regular tasks showed a difference between the A and B phase in the expected direction. The decline was not statistically significant (A: 24.4, B: 11.4, m: 18.9, $p = .14$). The change in non-specific prompts was statistically significant and in the right direction ($p = .048$, Figure 3 *online*). The change in corrections from trainer and specific prompts was not statistically significant and also not in the right direction ($p = .19$ and $p = .13$, respectively, Figure 1). The total amount of self-control behaviour remained stable (A: .9, B: 1.0, m: 1.0, $p = .62$), as did the subcategories self-corrections ($p = .86$) and requests for help ($p = .48$).

Statistical analyses of the performance on the repetitive tasks (Figure 2) did not show a significant change. Visual analyses showed, contrary to expectation, an increase in feedback (A: 7.3, B: 16.1, m: 27.8, $p = .65$). The changes in the level of the subcategories were also not statistically significant (non-specific prompts $p = .29$, specific prompts p

Table 2. Patients' characteristics.

Patient	Sex	Age	ABI	Days post-onset	Cognitive deficits (>2 SD below mean on NPA)	Behaviour	PCRS d.s.	SRSI
S	f	68	Ischemic CVA right-side	56	Memory, mental speed, executive functioning	Indifferent, flattened affect, poor initiative	45	50
K	m	55	Haemorrhage right thalamus	45	Memory, mental speed, executive functioning	Friendly, stress avoiding, quiet	18	50
R	m	60	Ischemic CVA right-side	66	Attention (sustained, neglect), executive functioning	Easily offended, inpatient, rebellious	40	50
B	m	57	TBI right-side due to accident while repairing motor cycle	30	Visual attention, mental speed, visual memory, executive functioning	Euphoric, helpful, verbally disinhibited	62	50

Information about patients S, K, R and B regarding sex, age, ABI, days post-onset, cognitive deficits as measured on the neuropsychological assessment (NPA), short description of behaviour, the difference score (d.s.) on the PCRS (Fordyce & Roueche, 1986), calculated by subtracting the scores of the significant other to the scores of the patient, and the score on the SRSI (Ownsworth et al., 2000), f = female, m = male, SD = standard deviation.

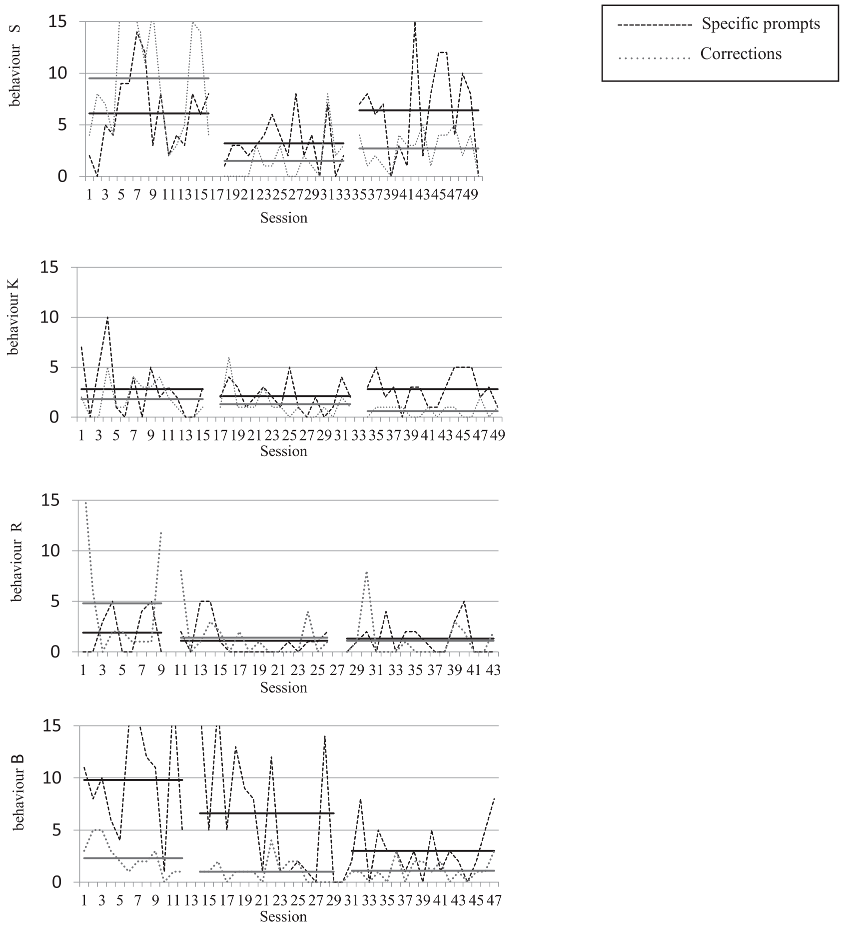

Figure 1. Performance on regular tasks: patients S, K, R, B.

Note: Results of patients S, K, R and B regarding the frequency of specific prompts and corrections per 10 minutes of regular tasks in the baseline, intervention and follow up phases. Patient S: data points 1–16 represent A1 sessions, data points 17–33 represent B sessions, and data points 34–50 represent the A2 sessions. Patient K: 1–15 A1, 16–32 B, 33–49 A2. Patient R: 1–9 A1, 10–26 B, 27–43 A2. Patient B: 1–13 A1, 14–30 B, 31–46 A2.

= .86 and trainer corrections $p = .52$, Figure 3, Figure 4 *online*). The total rate of self-control behaviour did not statistically significantly change (A: .6, B: 1.1, m: 3.9, $p = .73$). Visual inspection showed a decline, contrary to expectation. The levels within the sub-categories self-corrections ($p = .66$) and requests for help ($p = .76$) did not differ significantly either.

Randomisation qualified patient K to 15 baseline measurements, and therefore he completed a total of 47 training sessions (A: 15, B: 16, m: 16) spread over 18 weeks. Tasks were related to K's job as a municipal official (e.g., performing calculations) and his hobbies (e.g., maintaining the garden, visiting friends). The total amount of feedback on the regular tasks did not show statistically significant change (A: 9.8, B: 7.1, m: 9.3, $p = .57$), nor did the corrections made by the trainer ($p = .09$). Visual inspection showed a decline that conformed to expectation. Additionally, changes in specific prompts and

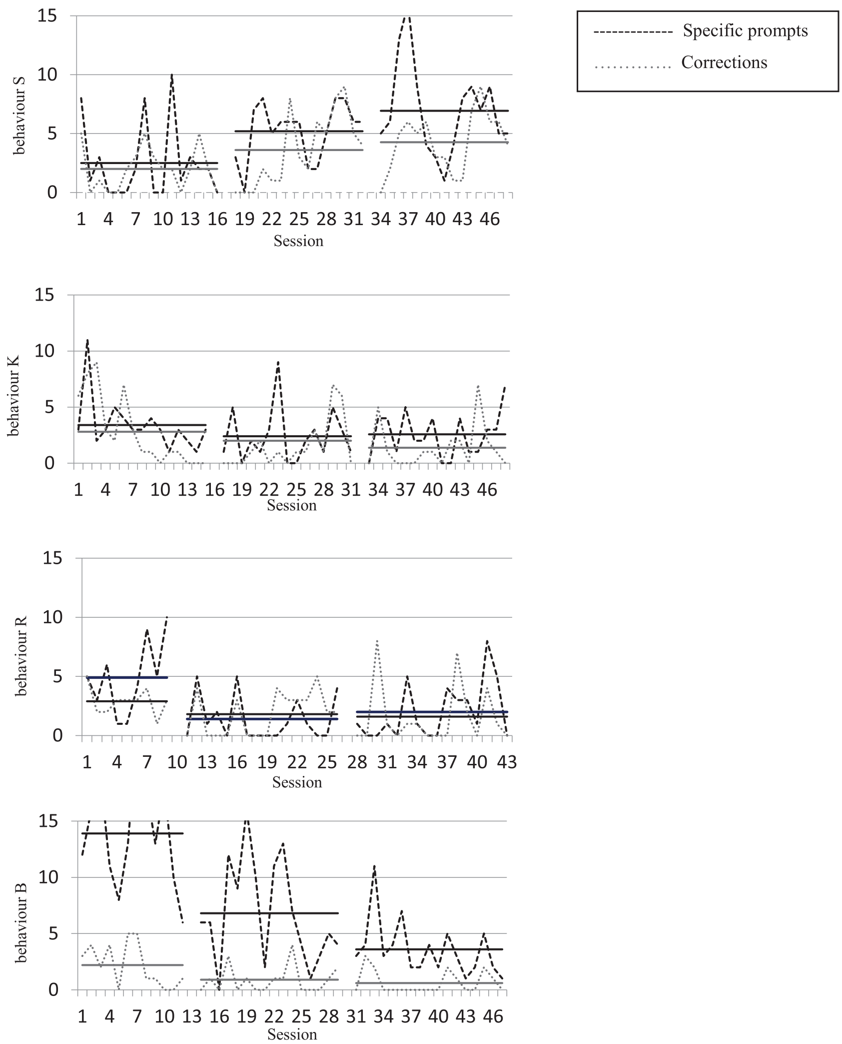

Figure 2. Performance on repetitive tasks: patients S, K, R, B.

Note: Results of patients S, K, R and B regarding the frequency of specific prompts and corrections per 10 minutes of regular tasks in the baseline, intervention and follow up phases. Patient S: data points 1–16 represent A1 sessions, data points 17–33 represent B sessions, and data points 34–50 represent the A2 sessions. Patient K: 1–15 A1, 16–32 B, 33–49 A2. Patient R: 1–9 A1, 10–26 B, 27–43 A2. Patient B: 1–13 A1, 14–30 B, 31–46 A2.

non-specific prompts were not statically significant (non-specific prompts, $p = .38$, specific prompts, $p = .43$). Visual inspection showed differences in the expected direction. The changes in observable self-control behaviour, specifically requests for help and self-corrections, were not statistically significant (respectively $p = .35$ and $p = .40$). Visual analyses showed an increase in the domain requests for help, which is in the expected direction, and a visual decrease in the domain self-corrections, which is not in the expected direction.

The total feedback score on the repetitive tasks did not show statistically significant change (A: 12.5, B: 8.8, m: 9.1, $p = .33$). Visual inspection showed a decrease, which

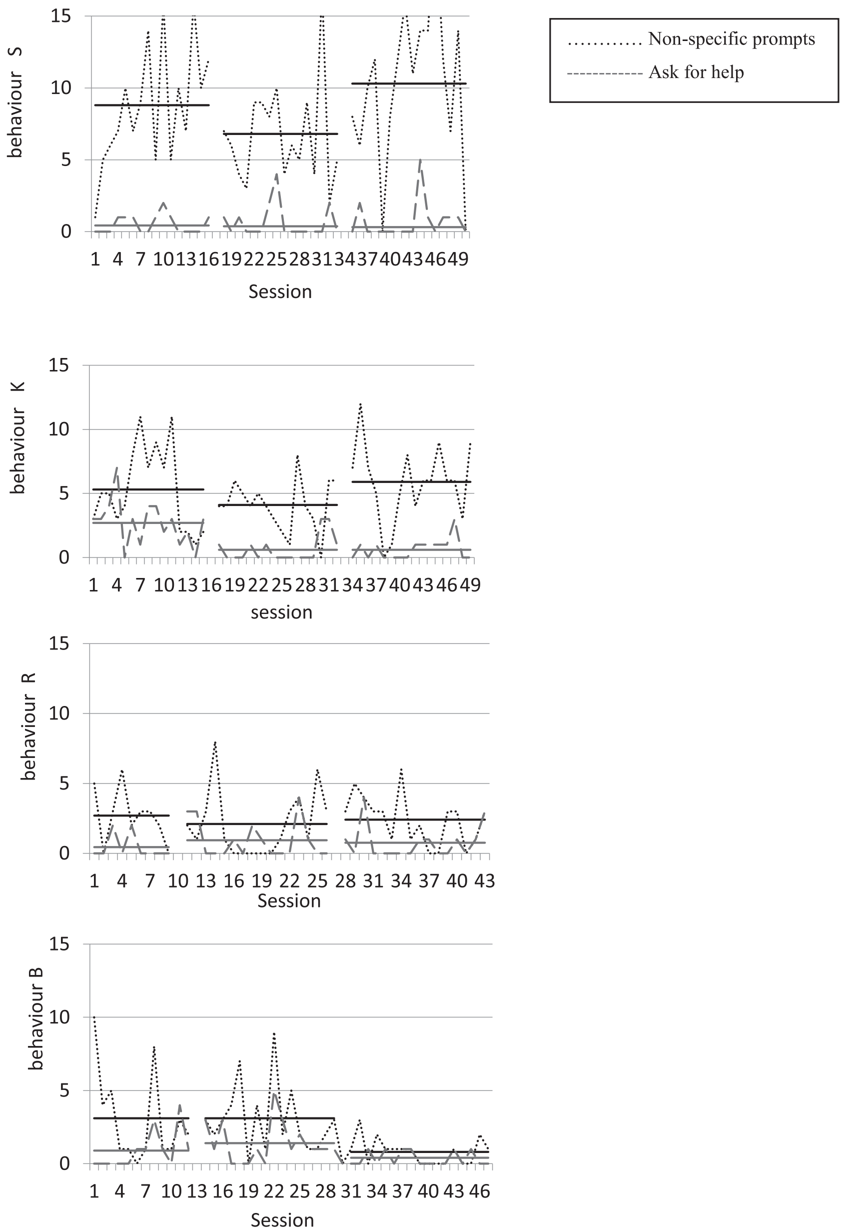

Figure 3. Performance on regular tasks: patients S, K, R, B.

Note: Results of patients S, K, R and B regarding the frequency of non-specific prompts and asking for help per 10 minutes of regular tasks. Patient S: data points 1–16 represent A1 sessions, data points 17–33 represent B sessions, and data points 34–50 represent A2 sessions. Patient K: 1–15 A1, 16–32 B, 33–49 A2. Patient R: 1–9 A1, 10–26 B, 27–43 A2. Patient B: 1–13 A, 14–30 B, 31–46 A2.

conformed to expectation. None of the three feedback subcategories showed a statistically significant change (non-specific prompts, $p = .33$, specific prompts, $p = .43$, trainer corrections, $p = .48$). They all decreased visually, which conformed to expectation. Self-

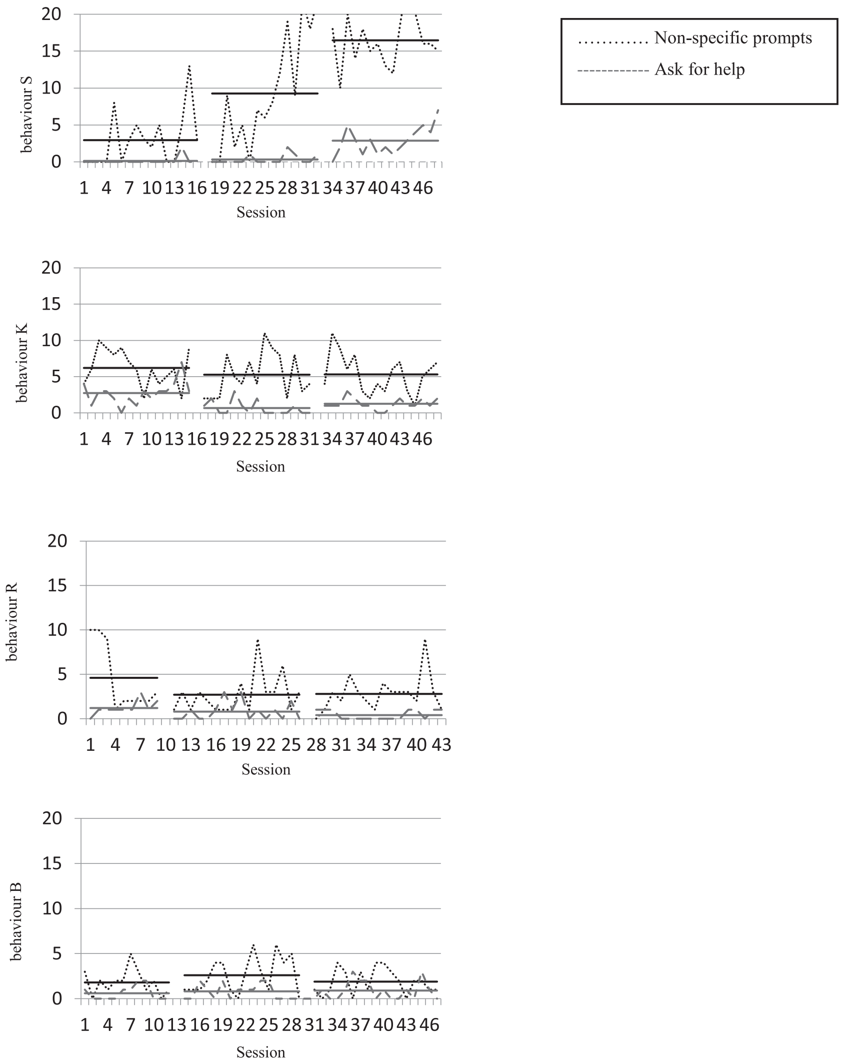

Figure 4. Performance on repetitive tasks: patients S, K, R, B.

Note: Results of patients S, K, R and B with respect to frequency of non-specific prompts and asking for help per 10 minutes of performing a task-specific learning task. Patient S: data points 1–16 represent phase A sessions, data point 17–33 represent phase B sessions, and data points 34–50 represent follow-up phase sessions. Patient K: 1–15 phase A, 16–32 phase B, 33–49 FU phase. Patient R: 1–9 phase A, 10–26 phase B, 27–43 FU phase. Patient B: 1–13 phase A, 14–30 phase B, 31–47 FU phase.

control behaviour (self-corrections and asking for help in total) did not show statistically significant change (A: 13, B: 9, m: 9, $p = .33$). It showed a visual decline, contrary to expectation. The decline in asking for help was statistically significant ($p = .048$) and in not the expected direction. The amount of requests for help did not change statistically significantly ($p = .14$), nor visually.

Randomisation qualified patient R to 9 baseline measurements, with a total of 41 sessions (A: 9, B: 16, FU: 16) spread over 13 weeks. Tasks were related to his work as a financial controller (verifying data, planning and control, financial administration) and to cooking and cycling. On the regular tasks, the change in total amount of feedback was not statistically significant (A: 9.3, B: 4.6, m: 4.8, $p = .09$). Visual inspection showed a change in the expected direction. The subcategory corrections needed from the trainer showed a statistically significant change ($p = .048$) and visually in the right direction. The rate of specific and non-specific prompts did not change statistically significantly (respectively $p = .33$, $p = .29$), but changed visually in the expected direction. There was no statistically significant change, nor visual change in the rate of self-control (A: 1.8, B: 1.7, m: 1.1, $p = .62$), self-corrections ($p = .43$) and requests for help ($p = .52$).

The total feedback score on the repetitive tasks showed a statistically significant decline in the rate of feedback overall (A: 12.3, B: 5.8, m: 5.9, $p = .048$) and this change was visually in the right direction. The scores on the three subcategories were not statistically significant (specific prompts $p = .09$; non-specific prompts $p = .14$; corrections by trainer $p = .20$), but visually showed changes in the expected direction on these three subcategories. The rate of self-control behaviour did not change statistically significantly (A: 2.2, B: .9, m: .8, $p = .12$), although visual analyses showed changes in the expected direction. The rate of self-corrections changed statistically significantly ($p = .048$) and changed visually, conforming with expectation. The rate of requests did not change statistically ($p = .57$) nor visually.

Randomisation qualified patient B to 13 baseline measurements, with a total of 45 training sessions (A: 13, B: 16, m: 16) spread over 15 weeks. The tasks were related to B's work as a truck driver (map reading, route planning, comparing and checking lists), as an organiser of foreign tours (making websites, checking websites of competitors), and to motorcycling (finding/buying components on the internet). The rate of feedback during regular tasks declined statistically significantly (A: 15.2, B: 10.7, m: 4.9, $p = .048$) and also visually in the expected direction. The change in feedback was largely attributed to a decrease in specific prompts ($p = .048$). The amount of corrections by the trainer did not change statistically significantly ($p = .20$), but visually showed a change in the expected direction. There was no statistically significant change, nor visual change, in the expected direction in non-specific prompts ($p = .38$). The rate of observable self-control behaviour remained stable across phases (A: 1.3, B: 1.5, m: .7, $p = .52$), as did the rate of self-corrections ($p = .29$) and requests for help ($p = .62$).

The rate of feedback on the repetitive tasks did not show a significant change ($p = .28$). Visual inspection showed change in the expected direction (A: 18.7, B: 11.5, m: 7.4). The change in specific prompts was not statistically significant ($p = .19$), although visual analyses showed a change in the expected direction. Non-specific prompts and corrections from the trainer did not change statistically significantly (respectively $p = .19$ and $p = .24$). Visual inspection showed a change in the expected direction. The rate of self-control behaviour did not change statistically significantly (A: .8 B: 1.2 m: 1.2, $p = .09$). The visual increase conformed with expectation. The change in request for help was not statistically significant ($p = .09$). The change was visually in the expected direction. There was no statistically significant change in amount of self-corrections ($p = .43$) and this was visually not in the expected direction.

Secondary outcome: General awareness, emotional distress

Patient S: Visual analyses showed that the PCRS difference score fluctuated without clear improvement (A: 42–45–45–52, B: 52–50–42–43, m: 47–42–43–45). The effect size (NAP) was weak (.5). The SRSI scores remained stable during baseline (A: 50–50–50–50), were relatively stable during the intervention (B: 50–50–45–50) and changed in the expected direction (m: 44–42–42–44). NAP showed a medium effect (.8). Visual analyses showed that the depression and anxiety levels, measured every 2 weeks, represented no symptoms (HADS total A:2–5 B: 3-1 m:6–6).

Patient K: The PCRS difference scores showed much variation overall (A: 14–18–8–14–18–12, B: –1–22–11–23, m: 8–13–22–20). They did not change in the expected direction, which was confirmed by a weak NAP score of .34. The SRSI scores changed in the expected direction and then remained relatively stable in the maintenance (A: 50–50–50–50–50–50, B: 50–50–27–26, m: 26–22–25–25) (NAP score .86, medium effect). The depression and anxiety levels showed moderate to severe symptoms during the baseline phase, mild to no symptoms during the intervention phase, and no symptoms during the maintenance phase (HADS total A: 17–19–12, B: 9–4, m: 6–5).

Patient R: There was a change in the expected direction in the PCRS difference scores (A: 40–28–32–35, B: 38–35–28–26, m: 28–28–22–27), corresponding to a medium effect (NAP .78). The SRSI curve also showed a decline (A: 47–41–47–44, B: 45–50–28–26, m: 32–22–10–5), which corresponded to a medium effect (NAP .83). Anxiety and depression levels represented no symptoms and remained stable (HADS total A: 2–5, B: 1–3, m:6–6).

Patient B: The PCRS difference scores and the SRSI score both showed a change in the expected direction (PCRS A: 63–56–64–60–56–54, B: 52–50–48–48, m: 46-47–44–42; SRSI A: 50–42–47–38–27–31, B: 29–32–30–23, m: 14–15–18–15). NAP showed a strong effect for both the PCRS (NAP 1.0) and the SRSI (NAP .95). B scored zero on the HADS self-report, which indicated that no symptoms were present (A: 0–0–0, B: 0–0, m: 0–0). Table 3 shows a summary of the results for the four patients in the study

Feasibility

All four participants completed the questionnaire. The entire programme (baseline, intervention, maintenance) was assessed to be good (Table 4), with a median of 9 out of 10. The duration and intensity of the therapy was rated to be good, as was applicability to daily life. All patients stated that the goal of the treatment was clear and that they were satisfied with participating in the study. The applicability to daily life was judged to be good by all patients, as was the duration and the total number of treatment sessions. The application of the treatment protocol was rated by the trainer to be fairly easy, but only after a fair amount of practice.

Discussion

The effectiveness and feasibility of a 4-week behavioural intervention method to improve awareness of deficits was evaluated in four patients with ABI, based on visual analyses whether changes are in the desired direction and based on statistical analyses (randomisation test and NAP). The results of the four SCEDs showed that on the regular tasks all patients needed less feedback from the trainer. This change was statistically significant in 3 out of 4 patients (fewer non-specific and specific prompts, fewer corrections). As the regular tasks differ in content throughout the training, the decline in trainer feedback reflects that there was generalisation across different tasks, which can be regarded as an improvement in the self-regulative aspect of

Table 3. Summary of results.

		S Visible change in the right direction	Statistically significant ($p \leq 0.05$)	K Visible change in the right direction	Statistically significant ($p \leq 0.05$)	R Visible change in the right direction	Statistically significant ($p \leq 0.05$)	B Visible change in the right direction	Statistically significant ($p \leq 0.05$)
Repetitive tasks	Total rate of feedback	no	no	yes	no	yes	yes	yes	no
	Non-specific prompts	no	no	no	no	yes	no	no	no
	Specific prompts	no	no	yes	no	yes	no	yes	no
	Trainer corrections	no	no	yes	no	yes	no	no	no
	Total rate of self-control	no	no	no	no	yes	no	no	no
	Self-corrections	no	no	no	.no	yes	yes	yes	no
	Requests for help	no	no	yes	yes	no	no	yes	no
Regular tasks	Total rate of feedback	yes	no	yes	no	yes	no	yes	yes
	Non-specific prompts	yes	yes	yes	no	yes	no	no	no
	Specific prompts	yes	no	yes	no	yes	no	yes	yes
	Trainer corrections	yes	no	yes	no	yes	yes	yes	no
	Total rate of self-control	no	no	no	no	no	no	no	no
	Self-corrections	no	no	yes	no	no	no	no	no
	Requests for help	no	no	no	no	no	no	no	.no
PCRS		no	NAP .50	no	NAP .34	yes	NAP .78	yes	NAP 1.0
SRSI		yes	NAP .80	yes	NAP .86	yes	NAP .83	yes	NAP .95

Overview of the improvements of patients S, K, R and B on the outcome measures feedback, control behaviour, the PCRS (Fordyce & Roueche, 1986) and the SRSI (Ownsworth et al., 2000); a "no" reflects no improvement, and a "yes" sign indicates a significant improvement ($p \leq 0.05$). Visible change in the right direction was classified as "yes" in case of a visual decrease in rate of feedback or a visual increase in self-control. NAP = nonoverlap of all pairs (Parker & Vannest, 2009).

Table 4. Feasibility of the Socratic feedback intervention.

	Ranking the quality (0–10)	Applicability to daily life	Patients' remarks
S	9	good	"It does not go like before. It needs to go slower, and I need to ask more. I make more mistakes, they say. I can think better now. I know where I am good at and where I have problems with."
K	8	good	"I have problems with memory and concentration. I forget to offer guests a drink, and I forget what I read in newspapers. I should work slower in order not to make too many mistakes, and I need to read the newspaper twice a day. I know my problems and I am able to use strategies to cope with them. I have learned to do things at a slower pace, to remain calm when a task is more difficult than I thought, and to think of strategies before starting a task."
R	9	good	"I have problems with time orientation, planning and concentration, which lead to problems at work and at home. It bothers me and makes me insecure. I am thinking about changing my workload and tasks at work. I am planning to delegate more to others, to depend more on my wife, and accept this, and to look frequently at the clock and in his schedule. I need to do one thing at a time and take into account my neglect."
B	10	good	"I progressed a lot. I have to stay concentrated and stay calm. Planning first before action. The training motivated me to explore the limits of cognitive functioning at home and stimulated me to talk about cognition with my wife."

Evaluation of the quality and applicability of the programme by patients, including quotes that were sampled on the last SRSI (patients) and on the feasibility questionnaire (patients).

awareness. Only one patient showed visually an increase in overt self-correction, although this change was statistically not significant. On the repetitive tasks, 3 out of 4 patients showed visual improvements in self-control behaviour (asked fewer questions, fewer self-corrections); in 2 out of 4 patients, these improvements were significant. As this improvement in self-control behaviour was not shown on the regular tasks, it is more likely that the improvement in self-control behaviour reflects task-specific learning instead of an improvement in awareness.

Awareness of deficits improved in all four patients on the SRSI, and in 2 out of 4 patients on the PCRS. Emotional distress was present in 1 out of 4 patients during the baseline phase and in none of the patients in the maintenance phase. With regard to the feasibility of the awareness intervention, patients and trainers rated the therapy to be feasible. The duration and intensity of the therapy was rated to be good, as was the applicability to daily life. The application of the treatment protocol was rated by the trainer to be fairly easy, but only after a fair amount of practice. These findings modestly suggest that non-confrontational feedback following the Socratic method might be a promising method for improving the self-awareness of deficits in ABI patients.

Our finding that a non-confrontational awareness programme can lead to improvements in awareness in terms of decreases in trainer-feedback is in agreement with previous research (Goverover et al., 2007; Ownsworth et al., 2006; Rebmann & Hannon, 1995; Tham, Ginsburg, Fisher, & Tegner, 2001). A decrease in specific trainer prompts was found in a study on the effects of non-confrontational awareness training (Ownsworth et al., 2006) derived from the "pause, prompt and praise" technique (Glynn et al., 1979). Our study did not replicate these findings with respect to an improvement in self-correction. A possible explanation for this is that observations of self-corrections is more difficult to objectify than trainer prompts and corrections: when the patient makes an error, the trainer pauses to give the patient the opportunity to correct the mistake. If self-correction does not occur, the trainer gives a non-specific prompt. It is possible that the trainers in our study paused a shorter period of time than the observers in the Ownsworth study.

The findings with respect to the improvement on the SRSI are in line with previous research (Goverover et al., 2007; Toglia et al., 2010). The PCRS improved in 2 out of 4 patients. Differences in responses to general versus specific questionnaires have been observed previously (Sherer et al., 1998). The discrepancies could reflect the differences in psychometric properties of the instruments used (Smeets, Ponds, Verhey, & van Heugten, 2012). For example, the PCRS provides restricted possibilities to answer whereas the SRSI has an interview style with open-ended questions. We observed that on the PCRS, patients with executive problems tend to answer quickly, impulsively and incompletely. Furthermore, the PCRS and SRSI measure different aspects of an awareness of deficits (Smeets et al., 2012).

We checked the study's internal and external validity with the RoBiNT scale. The total scale score was 22 out of 30, the internal validity subscale was 7/14, the external validity and interpretation subscale was 15/16 (Tate et al., 2013). See Appendix 4 for scoring per item. In terms of the internal validity, the elements randomisation, sampling, blind assessor and inter-rater reliability can be scored positive and the elements design, blinding of the therapists and patients and treatment adherence score negative. Concerning the external validity, it must be noted that the generalisability of the results may be somewhat limited because all patients were treated in the same setting by the same treatment team.

This study also had several limitations that should be addressed. During the analysis of the results, it became apparent that the measurement of self-control behaviour was inadequate. It is insufficient to only measure self-correcting behaviour spoken out loud, as patients can correct themselves without verbalising. Hence, the results of the self-correction had no additional value. A second limitation is that the generalisability to other contexts and other patients is restricted. A third limitation concerns the feedback method of the maintenance phase. It is very difficult for a cognitive trainer to "unlearn" a feedback method such as Socratic feedback. Although the trainer stated that he was able to give feedback in the original way by reading the protocol before every session, some interference might have occurred. Our evaluation of self-regulatory behaviour (amount of feedback needed, error self-correction and requests for help) was based on the client's performance when they were with the cognitive trainer, who was providing them with corrective feedback during the baseline and maintenance sessions and Socratic feedback during the treatment sessions. Self-regulatory behaviour therefore may have been different had we evaluated the client's task performance separately from their trainers (i.e., performing the task on their own or with an assessor who used a uniform approach to prompting across all three phases). Fourth, with only 21 randomisations, statistical significance at a $p \leq .05$ level can only be achieved if the actual difference is the largest of all possible 21 differences, which might imply that the statistical test is not powerful enough to detect potentially relevant differences between phases (Ferron & Sentovich, 2012). With respect to feedback during the regular tasks we found that although only 4 out of 16 changes were statistically significant, nearly all changes (15 out of 16) were in the right direction and herewith the visual analyses might add information to the effectiveness of the intervention. Fifth, the single-case reporting guideline in behavioural interventions (SCRIBE) (Tate et al., 2016) helped us to prepare the reports of this single-case research in terms of clarity, completeness, accuracy and transparency. Treatment and procedural fidelity were not assessed formally in this study. Last, patients in the post-acute phase are likely to show some spontaneous recovery of cognitive dysfunctions. The use of randomisation is a means to control for this influence.

In this study, we have shown that the method of non-confrontational feedback following the Socratic feedback method, which is a key element in CBT, might be a promising tool to improve the self-awareness of deficits in ABI patients. Further research is needed, preferably a randomised controlled trial, to obtain additional evidence about the effectiveness of this new intervention method.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Allen, C. C., & Ruff, R. M. (1990). Self-rating versus neuropsychological performance of moderate versus severe head-injured patients. *Brain Injury*, *4*(1), 7–17.

Anson, K., & Ponsford, J. (2006). Coping and emotional adjustment following traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *21*(3), 248–259.

Bach, L. J., & David, A. S. (2006). Self-awareness after acquired and traumatic brain injury. *Neuropsychological Rehabilitation*, *16*(4), 397–414.

Barco, P. P., Crosson, B., Bolesta, M. M., Werts, D., & Stout, R. (1991). Training awareness and compensation in postacute head injury rehabilitation. In J. S. Kreutzer & P. H. Wehman (Eds.), *Cognitive rehabilitation for persons with traumatic brain injury: A functional approach* (pp. 129–146). Baltimore, MD: Paul H. Brookes Publishing.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. Philadephia: University of Pennsylvania Press.

Bogod, N. M., Mateer, C. A., & Macdonald, S. W. S. (2003). Self-awareness after traumatic brain injury: A comparison of measures and their relationship to executive functions. *Journal of the International Neuropsychological Society*, *9*(3), 450–458.

Bransford, J. D., Brown, A. L., Cocking, R. R., Donovan, M. S., & Pellegrino, J. W. (2000). *How people learn: Brain, mind, experience, and school (expanded edition)*. Washington, DC: National Academy Press.

Bulté, I., & Onghena, P. (2008). An R package for single-case randomization tests. *Behavior Research Methods*, *40*(2), 467–478.

Bulté, I., & Onghena, P. (2012). When the truth hits you between the eyes: A software tool for the visual analysis of single-case experimental data. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *8*(3), 104–114.

Calero-Elvira, A., Froján-Parga, M. X., Ruiz-Sancho, E. M., & Alpañés-Freitag, M. (2013). Descriptive study of the Socratic method: Evidence for verbal shaping. *Behavior Therapy*, *44*(4), 625–638.

Cicchetti, D. V. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clininical and Experimental Neuropsychology*, *23*(5), 695–700.

Crosson, B., Barco, P. P., Velozo, C. A., Bolesta, M. M., Cooper, P. V., Werts, D., & Brobeck, T. C. (1989). Awareness and compensation in postacute head injury rehabilitation. *Journal of Head Trauma Rehabilitation*, *4*(3), 46–54.

Edgington, E., & Onghena, P. (2007). *Randomization tests*. Boca Raton, FL: CRC Press.

Edgington, E. S. (1980). Random assignments and statistical tests for one-subject experiments. *Behavioral Assessment*, *2*, 19–28.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York, NY: Lile Stuart.

Ferron, J., & Sentovich, C. (2002). Statistical power of randomization tests used with multiple-baseline designs. *The Journal of Experimental Education*, *70*(2), 165–178.

Flashman, L. A., & McAllister, T. W. (2002). Lack of awareness and its impact in traumatic brain injury. *NeuroRehabilitation*, *17*(4), 285–296.
Fleming, J. M., Strong, J., & Ashton, R. (1996). Self-awareness of deficits in adults with traumatic brain injury: How best to measure? *Brain Injury*, *10*(1), 1–16.
Fordyce, D. J., & Roueche, J. R. (1986). Changes in perspectives of disability among patients, staff, and relatives during rehabilitation of brain injury. *Rehabilitation Psychology*, *31*(4), 217–229.
Giacino, J. T., & Cicerone, K. D. (1998). Varieties of deficit unawareness after brain injury. *The Journal of Head Trauma Rehabilitation*, *13*(5), 1–15.
Glynn, T., McNaughton, S., Robinson, V., & Quinn, M. (1979). *Remedial reading at home: Helping you to help your child*. Wellington: New Zealand Council for Educational Research.
Goverover, Y., Johnston, M. V., Toglia, J., & DeLuca, J. (2007). Treatment to improve self-awareness in persons with acquired brain injury. *Brain Injury*, *21*(9), 913–923.
Hart, T., Giovannetti, T., Montgomery, M. W., & Schwartz, M. F. (1998). Awareness of errors in naturalistic action after traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *13*(5), 16–28.
Hart, T., Seignourel, P. J., & Sherer, M. (2009). A longitudinal study of awareness of deficit after moderate to severe traumatic brain injury. *Neuropsychological Rehabilitation*, *19*(2), 161–176.
Heyvaert, M., & Onghena, P. (2013). Analysis of single-case data: Randomisation tests for measures of effect size. *Neuropsychological Rehabilitation*, *24*(3–4), 507–527.
Kelley, E., Sullivan, C., Loughlin, J. K., Hutson, L., Dahdah, M. N., Long, M. K., & Poole, J. H. (2014). Self-awareness and neurobehavioral outcomes, 5 years or more after moderate to severe brain injury. *The Journal of Head Trauma Rehabilitation*, *29*(2), 147–152.
Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods*, *15*(2), 124–144.
Langer, K. G., & Padrone, F. J. (1992). Psychotherapeutic treatment of awareness in acute rehabilitation of traumatic brain injury. *Neuropsychological Rehabilitation*, *2*(1), 59–70.
Ledford, J. R., & Gast, D. L. (2014). Measuring procedural fidelity in behavioural research. *Neuropsychological Rehabilitation*, *24*, 332–348.
McNaughton, S. T., Glynn, T., & Robinson, V. (1987). *Pause prompt and praise: Effective tutoring of remedial reading*. Birmingham, England: Positive Products.
Neenan, M. (2009). Using Socratic questioning in coaching. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *27*, 249–264.
Olssøn, I., Mykletun, A., & Dahl, A. A. (2005). The hospital anxiety and depression rating scale: A cross-sectional study of psychometrics and case finding abilities in general practice. *BMC Psychiatry*, *5*, 46.
Onghena, P. (1992). Randomization tests for extensions and variations of ABAB single-case experimental designs: A rejoinder. *Behavioral Assessment*, *14*, 153–153.
Overholser, J. C. (1993). Elements of the Socratic method: I. Systematic questioning. *Psychotherapy*, *30*(1), 67–74.
Ownsworth, T., Fleming, J., Desbois, J., Strong, J., & Kuipers, P. (2006). A metacognitive contextual intervention to enhance error awareness and functional outcome following traumatic brain injury: A single-case experimental design. *Journal of the International Neuropsychological Society*, *12*(1), 54–63.
Ownsworth, T., Quinn, H., Fleming, J., Kendall, M., & Shum, D. (2010). Error self-regulation following traumatic brain injury: A single case study evaluation of metacognitive skills training and behavioural practice interventions. *Neuropsychological Rehabilitation*, *20*(1), 59–80.
Ownsworth, T. L., McFarland, K., & Young, R. M. (2000). Development and standardization of the self-regulation skills interview (SRSI): A new clinical assessment tool for acquired brain injury. *The Clinical Neuropsychologist*, *14*(1), 76–92.
Padesky, C. (1993). *Socratic questioning: Changing minds or guided discovery?* Keynote address delivered at the European Congress of Behavioural and Cognitive Therapies, London.
Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*(4), 357–367.
Perdices, M., & Tate, R. L. (2009). Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognised and undervalued? *Neuropsychological Rehabilitation*, *19*(6), 904–927.
Port, A., Willmott, C., & Charlton, J. (2002). Self-awareness following traumatic brain injury and implications for rehabilitation. *Brain Injury*, *16*(4), 277–289.
Pouwer, F., Van Der Ploeg, H. M., Ader, H. J., Heine, R. J., & Snoek, F. J. (1999).). The 12-item well-being questionnaire. An evaluation of its validity and reliability in Dutch people with diabetes. *Diabetes Care*, *22*(12), 2004–2010.

Prigatano, G. P., Altman, I. M., & O'brien, K. P. (1990). Behavioral limitations that traumatic-brain-injured patients tend to underestimate. *The Clinical Neuropsychologist*, *4*(2), 163–176.

Rebmann, M. J., & Hannon, R. (1995). Treatment of unawareness of memory deficits in adults with brain injury: Three case studies. *Rehabilitation Psychology*, *40*(4), 279–287.

Roth, A. D., & Pilling, S. (2007). *Competences required to deliver effective cognitive and behavioural therapy for people with depression and with anxiety disorders*. London: HMSO, Department of Health. 43, 673–680.

Schmidt, J., Lannin, N., Fleming, J., & Ownsworth, T. (2011). Feedback interventions for impaired self-awareness following brain injury: A systematic review. *Journal of Rehabilitation Medicine*, *43*(8), 673–680(8).

Schrijnemaekers, A.-C. M. C., Smeets, S. M., Ponds, R. W., van Heugten, C. M., & Rasquin, S. (2014). Treatment of unawareness of deficits in patients with acquired brain injury: A systematic review. *The Journal of Head Trauma Rehabilitation*, *29*(5), 9–30.

Sherer, M., Bergloff, P., Levin, E., High Jr, W. M., Oden, K. E., & Nick, T. G. (1998). Impaired awareness and employment outcome after traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *13*(5), 52–61.

Sherer, M., Hart, T., & Nick, T. G. (2003). Measurement of impaired self-awareness after traumatic brain injury: A comparison of the patient competency rating scale and the awareness questionnaire. *Brain Injury*, *17*(1), 25–37.

Smeets, S., Ponds, R., Verhey, F., & van Heugten, C. (2012). Psychometric properties and feasibility of instruments used to assess awareness of deficits after acquired brain injury: A systematic review. *The Journal of Head Trauma Rehabilitation*, *27*(6), 433–442.

Sohlberg, M. M., & Mateer, C. A. (2001). *Cognitive rehabilitation: An integrative neuropsychological approach*. New York, NY: Guilford.

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., & Wilson, B. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology*, *4*, 1–9.

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item risk of bias in N-of-1 trials (RoBiNT) scale. *Neuropsychological Rehabilitation*, *23*(5), 619–638.

Tham, K., Ginsburg, E., Fisher, A. G., & Tegner, R. (2001). Training to improve awareness of disabilities in clients with unilateral neglect. *The American Journal of Occupational Therapy*, *55*(1), 46–54.

Toglia, J., Johnston, M. V., Goverover, Y., & Dain, B. (2010). A multicontext approach to promoting transfer of strategy use and self regulation after brain injury: An exploratory study. *Brain Injury*, *24*(4), 664–677.

Trudel, T. M., Tryon, W. W., & Purdum, C. M. (1998). Awareness of disability and long-term outcome after traumatic brain injury. *Rehabilitation Psychology*, *43*(4), 267–281.

Urbaniak, G. C., & Plous, S. (2013). Research Randomizer (Version 4.0) [Computer software]. Retrieved from June 22, 2013.

Vanderploeg, R. D., Belanger, H. G., Duchnick, J. D., & Curtiss, G. (2007). Awareness problems following moderate to severe traumatic brain injury: Prevalence, assessment methods, and injury correlates. *Journal of Rehabilitation Research and Development*, *44*(7), 937–950.

World Health Organization. (2001). *International classification of functioning, disability and health*. Geneva: World Health Organization.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370.