Psychiatry Research 245 (2016) 74–82

Contents lists available at ScienceDirect

Psychiatry Research

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

Reaction time as an indicator of insufficient effort: Development and validation of an embedded performance validity parameter

Andreas Stevens[a,b,*], Simone Bahlo[a], Christina Licha[a], Benjamin Liske[a,b], Elisabeth Vossler-Thies[a]

[a] Medizinisches Begutachtungsinstitut Tübingen, Doblerstr. 17, D-72074 Tübingen, Germany
[b] University of Tübingen, Dept. of Psychiatry, Calwerstr. 14, D-72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 7 May 2016
Received in revised form 19 July 2016
Accepted 6 August 2016
Available online 8 August 2016

ABSTRACT

Subnormal performance in attention tasks may result from various sources including lack of effort. In this report, the derivation and validation of a performance validity parameter for reaction time is described, using a set of malingering-indices ("Slick-criteria"), and 3 independent samples of participants (total n =893). The Slick-criteria yield an estimate of the probability of malingering based on the presence of an external incentive, evidence from neuropsychological testing, from self-report and clinical data. In study (1) a validity parameter is derived using reaction time data of a sample, composed of inpatients with recent severe brain lesions not involved in litigation and of litigants with and without brain lesion. In study (2) the validity parameter is tested in an independent sample of litigants. In study (3) the parameter is applied to an independent sample comprising cooperative and non-cooperative testees. Logistic regression analysis led to a derived validity parameter based on median reaction time and standard deviation. It performed satisfactorily in studies (2) and (3) (study 2 sensitivity=0.94, specificity=1.00; study 3 sensitivity=0.79, specificity=0.87). The findings suggest that median reaction time and standard deviation may be used as indicators of negative response bias.

1. Introduction

Attention is a complex psychological construct. Kahneman (1973) reviewed the concepts for attention and summarized that attention may be understood as a mechanism for controlling the significance of stimuli and selecting them accordingly (Kahneman, 1973, page 2). Kahneman (1973) already pointed to an effort component in attention and argued that effort is modulated by subjective factors, such as perceived incentives, and by external factors such as task complexity. Van Zomeren and Brouwer (1994) proposed 2 × 2 qualities of attention: Intensity (phasic or tonic) and focus (selective or divided). Traditionally, reaction time (RT) paradigms are used to assess attention, recording both speed (RT in ms) and accuracy (no. of errors). Impairment of attention may result from quite heterogeneous causes, such as decreased information processing capacity, substance use, reduced ability to focus on relevant stimuli, but also from lack of effort, e.g. when no relevant incentive is perceived. Brain injury may reduce performance in attention tasks, either through slowed processing speed or through inability to focus, leading to longer RT and an increased error rate (Collins and Long, 1996; Rueckert and Grafman, 1996; Van Zomeren and Deelman, 1978). Several investigators have observed that differences between healthy controls and brain injured patients are more salient when complex tasks are used, thus increasing the cognitive load (Collins and Long, 1996; Kujala et al., 1994; Reicker et al., 2007; Tombaugh et al., 2007; Van Zomeren and Deelman, 1976, 1978). After substantial brain injury prolonged reaction times may persist for years (Tombaugh et al., 2007; Van Zomeren and Deelman, 1978). In contrast, no enduring deficits have been found in patients with mild traumatic brain injury (TBI) (Dikmen et al., 1986; Woods et al., 2015a, 2015b, see also the WHO collaborative center task force study on the prognosis of mild TBI, based on 428 studies; Carroll et al., 2004). The type of attention deficit varies with the location of brain damage: Alexander et al. (2005) observed that only lesions to the right superomedial frontal region are regularly associated with prolonged reaction times (increase of about 160 ms), while subjects with lesions to other frontal brain areas performed normally. Moreover, the groups did not differ with respect to response accuracy. Apart from brain lesions and substances, drowsiness and insufficient effort may also cause slowed RTs (Reicker, 2008; Steck et al., 2000; Strauss et al., 1994).

The detection of insufficient effort – as opposed to biologically

* Corresponding author at: Medizinisches Begutachtungsinstitut Tübingen, Doblerstr. 17, D-72074 Tübingen, Germany.
E-mail address: stevens@med-begutachtung.de (A. Stevens).

http://dx.doi.org/10.1016/j.psychres.2016.08.022
0165-1781/

impaired processing speed – is thus of crucial relevance in medicolegal assessments. However, most stand-alone performance validity tests are based on memory paradigms and may thus be less sensitive to detect suboptimal effort in attention tasks. There have been attempts to spot deliberate slowing of motor responses by recording brain evoked potentials during a RT task. Kottler et al. (2010) examined healthy volunteers in the conditions "just watch", "react as fast as possible" and "feign a slow reaction". Deliberate slowing of motor responses was associated with a delayed P3 and a delayed as well as deformed lateralized readiness potential (LRP). A similar approach was used by Vagnini (2007). In three groups of participants (normal volunteers instructed to perform well, normal volunteers instructed to malinger, and a clinical group with documented brain injury instructed to perform honestly) RT, SD and event related brain potentials were measured while they performed the TOMM (Test of Memory Malingering, Tombaugh, 1997). Vagnini observed that RT in the TOMM as well as frontal brain potentials related to the old/new effect achieved good classification (malingering/non-malingering) when TOMM accuracy was used as standard of validity. Ord et al. (2010) and Erdodi et al. (2014) have recently proposed several validity indicators for the CPT-II (Conners, 2004), based on response quality and RT standard error. In the Ord et al. study, Slick et al. (1999) criteria were used as "gold standard" while Erdodi et al. employed the WMT (Word Memory Test, a stand-alone performance validity test, Green, 2003) and a combination of several other validity measures. In both studies, response quality and reaction speed declined when the likelihood of malingering increased. Failing more than two of the validity indicators of the RT task was associated with a sensitivity of 0.50 and a specificity of 0.95 for malingering (Erdodi et al., 2014). Woods et al. (2015b) described a novel test for simple RT measurement and reported that experimental simulators produced RTs that were about double the RT found in brain-injured participants. In another study Woods et al. (2015a) evaluated the performance of controls, experimental malingerers and brain-injured patients in a visual choice RT task. Experimental simulators needed roughly double the time to react than both patients and controls. While task complexity increases the variance between healthy and brain-injured testees, it seems less useful to diagnose insufficient effort (Stevens and Merten, 2010).

This report describes the development and validation of an embedded validity parameter derived from RT data gathered in naturalistic settings. Only data from "real" testees were used, i.e. from patients and litigants instead of experimental simulators.

2. Methods

This section is organized in three parts. In study (1) a validity indicator is derived by comparing a clinical sample of patients with acute brain lesions and no known external incentive with compensation seekers with and without brain damage. In study (2) the validity parameter is cross-validated in an unrelated archival sample of testees. One group of this sample comprised litigants who performed far below normal limits in the RT task, but were advised after the test that their performance was non-credible, because it would be by far incompatible with driving a car. When repeating the task, they showed a medically unexplainable acceleration of RTs by more than 2 SDs. The second group comprised litigants who according to Slick et al. criteria were honest performers. In study (3) the new validity parameter was applied to an unrelated archival sample of compensation seekers whose participants are either assumed to be honest performers according to Slick et al. criteria or as probably malingering.

The study reports on data from 893 persons totally. Inclusion criteria in studies 1–3 were ability to yield informed consent, age between 18 and 65 years, morphological assessment of brain, German School degree. Exclusion criteria were impaired vision (vision with both eyes >0.5 at 35 cm reading distance, assessed with standard reading probes), neglect (assessed according to Keller and Grömminger, 1995), aphasia, paresis of the writing hand, severe dysexecutive disorder (unability to obey test instructions) and a psychiatric diagnosis. From all participants written informed consent was obtained. The local ethics committee had reviewed and consented to the study. The data were pseudonymized and entered into an SPSS 13® (SPSS Inc. Chicago) data sheet.

2.1. Participants

2.1.1. Study (1): derivation of a validity indicator from archival clinical and assessment data

The authors are indebted to the 286 participants who volunteered for the study. These included $n=61$ inpatients undergoing rehabilitation treatment in the medical rehabilitation center Hohenurach, Germany. All of them had been diagnosed with acute brain lesions (group "acute brain lesion") as evidenced by MRI-imaging. Table 1 details demographic data, the brain regions involved and the underlying diseases. Glasgow Coma score ratings and documented posttraumatic amnesia were not available for most of these patients, as only 11 (18%) had suffered a traumatic brain injury.

$N=96$ persons were assessed in the Tübingen Forensic Institute seeking compensation for brain lesions suffered a median 3 years prior to assessment ("compensation brain lesion"). The third group consisted of $n=129$ persons seeking compensation for cognitive impairment following an accident, however, there was no evidence of a brain lesion by CCT or MRI ("compensation no brain lesion"). Not all participants of the latter group were evaluated with MRI and although iron-sensitive sequence were included in all MRIs done, no special techniques such as Diffusion Tensor imaging were performed. Thus, some cases with brain lesions may have been misclassified as "no brain lesion". On the other hand, for each case, clinical records and EEG recordings were available, also, a thorough neurological investigation as well as an EEG recording were done on the day of the assessment. When there were clinical or electroencephalographic sings of damage to the central nervous system, the case was not included in the category "no brain lesion" and a radiological examination was ordered. In none of these cases there was documented posttraumatic amnesia, although some of them reported posttraumatic amnesia at the time of assessment. The three groups differ in some respects with regard to their demographic variables (Table 1). As well as the location of brain lesions.

2.1.2. Study (2): validation in an archival data sample comprising known authentic and non-authentic responders in the RT-task

This sample was composed of 2 subgroups. Group „malingered reaction time" were testees who performed below percentile 16 (age and gender-corrected) in the alertness task described below. After they had solved some other tests of a cognitive test battery they were informed that their reaction speed casts doubt on their ability to drive. (German law requires drivers to perform better than percentile 15 in various attention-related tasks and professional drivers must perform better than percentile 32). These testees were asked to repeat the alertness task. Of $n=165$ such individuals 90% improved their performance in the 2nd run (i.e., after the warning) and 87/165 (51%) improved more than by 2 SDs. In detail, the following differences were obtained when the median RT between the two runs were compared: (percentile/median RT-difference (ms)): 20/35, 40/53, 60/105, 70/137, 80/197.