

タブレット版注意機能スクリーニング検査作成の試み I

—再検査信頼性の検討—

八田 武俊 (岐阜医療科学大学 保健科学部, hatta@u-gifu-ms.ac.jp)

重森 健太 (関西福祉科学大学 保健医療学部, shigemori@tamametama.ac.jp)

木村 貴彦 (関西福祉科学大学 健康福祉学部, takimura@tamametama.ac.jp)

八田 武志 (関西福祉科学大学 健康福祉学部, hatta@tamateyama.ac.jp)

岩原 昭彦 (京都女子大学 発達教育学部, iwahara@kyoto-wu.ac.jp)

Attempt to create a digital attention screening test I:

Examination of test-retest reliability based on comparison with paper-based test version

Taketoshi Hatta (Department of Health Science, Gifu University of Medical Sciences, Japan)

Kenta Shigemori (Department of Rehabilitation Science, Kansai University of Welfare Sciences, Japan)

Takahiko Kimura (Department of Health Science, Kansai University of Welfare Sciences, Japan)

Takeshi Hatta (Department of Health Science, Kansai University of Welfare Sciences, Japan)

Akihiko Iwahara (Department of Education, Kyoto Women's University, Japan)

Abstract

The paper-version D-CAT (digit cancellation test) was developed in 2001 and distributed widely. Pap. D-CAT was developed to provide a highly practical and user-friendly assessment of various aspects of prefrontal cortex functioning, including information processing speed, the ability to focus attention, and executive functioning. As part of the attempt to develop a digital-based D-CAT, in addition to the paper-based, an attempt was made to examine the reliability of the prototype (Dig. D-CAT) by comparing it with paper-based D-CAT. The reliabilities of Pap. D-CAT and Dig. D-CAT were evaluated using a test-retest paradigm. The reliability coefficient was lower for Dig. D-CAT than for Pap. D-CAT. The reliability coefficient for Pap. D-CAT was sufficiently high, and that for Dig. D-CAT had a reasonable level of correlation. The musculoskeletal load of stylus pressure control on the tablet screen may be the cause of the difference, and further verification is necessary.

Key words

digital attention screening test, test-retest reliability, D-CAT, digit cancellation test, focused attention

1. はじめに

本研究の目的は、2001年に開発され、専用の用紙で実施される注意機能検査 (Digit-Cancellation-Test) のタブレット版 (以下 Dig. D-CAT) を作成するための基礎的検討である。具体的には、試作した Dig. D-CAT (Ver.4) のユーザビリティや信頼性、現在使用されている紙ベースの D-CAT (以下 Pap. D-CAT) との比較を行うことで、問題点の有無や課題を確認する。

D-CAT は、①特別な機械装置や器具を必要としないこと、②対象者への心的負担が過度にならないこと、③検査者が実施と結果の判定に特別な訓練を要しないこと、④検査結果の採点・評価が簡単であること、⑤基準値を有し、機能評価が科学的に裏付けられた方法で可能なことなど、八田 (2007) が指摘した注意検査の 5 要件を満たしており、注意機能の評価する簡便な検査が当時なかったために、普及しつつある。

Pap. D-CAT とは、検査対象者が印字された乱数列から、検者によって指定されたターゲットとなる数字を制限時間内に抹消するという課題である。このとき、ターゲッ

トの数字は 1～3 つのパターンがあり、ターゲット数が多いほど、課題の難易度は高く、注意機能に関する認知的負荷も高い。また、制限時間は 30 秒の場合も想定されているが、基本的に 60 秒とされている。検索した文字数 (Total Performance: TP)、正しい数字を抹消した文字数 (Total Target: TT)、抹消すべきであるが見落とした文字数 (Total Omission: TO)、ターゲットでないにもかかわらず抹消した文字数 (False Alarm: FA) や、指定される文字数が増加するにつれて増える作業記憶の負荷を疲労指標とし、注意機能を構成する①集中の焦点化、②注意の維持、③注意の選択の要素を評価するのが主な目的である (伊藤・吉崎・八田, 2001)。疲労度が検出評価できることも、マニュアルに年齢別・性別の基準値とともに記載されている (八田・伊藤・吉崎, 2014)。この検査を利用した研究発表では、全ての条件や指標を使うというよりも、研究目的に沿って選択的な使用が行われることが多い。

D-CAT の理論的背景は、神経心理学の臨床家の提唱する注意理論の検討から、Sohlberg & Matter (1989) による理論に準拠する。Sohlberg & Matter (1989) は、注意を構成する要素を階層的に捉え、低次の階層は高次階層の基盤になるとしている。階層は、下位から①注意の焦点化 (focused attention)、②注意の維持 (sustained attention)、③選択的注意 (selective attention)、④注意の切り換え

(alternating attention)、⑤注意の配分 (divided attention) である。前述したように、D-CAT はこの注意理論の①から③までの要素を評価する目的で作成された。

これまで、多くのパーソナリティ検査法や知能検査などは、専用の用紙や機械、器具を用いるものであったが、近年の電子工学技術の画期的な進歩によって、その多くは紙ベースからコンピュータ画面を介した検査へと移行する傾向が認められる。このような背景から、Pap. D-CAT についても検査用紙をタブレット画面での表示を想定し、タッチパネル方式で実施できないかとする企画が進められている。タブレット版を開発すれば、検査時に用紙の手配や筆記具の準備などが一切不要となり、どこでも、だれでもすぐに利用でき、何よりも、採点にかかる作業負担を削減できるという利点がある。現在、紙ベースの D-CAT では単純に指標となる数を数えることで採点結果を得ることができるが、大量の場合、膨大な時間を要することや意図しないミスも発生しやすいため、専用の読み取り機器を使って採点を行う OCR 版 D-CAT もある。この OCR 版であれば、一定の採点負担は減じられるが、電子化データにするためには、転記作業が必要であること、さらに専用の機器が必要である。何より、将来的にデジタル機器の利用に馴染んだ世代が D-CAT を利用するとしたら、実施者や非検査者ともに結果の処理をコンピュータ処理できるように設計されたタブレット版が紙ベースの検査に置き換わるのは当然といえるだろう。

しかしながら、タブレット画面での刺激材料の提示やタッチパネル画面への慣れは、若年者でさほど問題にならなくとも、中高年齢者を含む全ての年齢層で同じ親近性とは考えにくい。コンピュータのない時代に育ってきた現在の中高齢者は、スマートフォンに代表される IT 技術が溢れる生活場面での戸惑いは小さくないし、熟練度も低い。そうした現状は、高齢者を対象とする IT 技術を用いた詐欺犯罪の発生にも表れている。つまり、課題自体を紙ベースからデジタル版に置き換えることは簡単でも、その測定内容や精度が同等であることを担保できるかということ、それほど単純な問題ではない。例えば、過去の紙ベースでの検査データとの比較利用を考えねばならない場面での変換係数の準備など、慎重に検討すべき問題を含んでおり、こうした検討すべき課題については順次取り組む必要がある。そこで、本研究ではまず、Pap. D-CAT のように乱数列をタブレット画面で表示し、タッチペンで指定された数字に触れるという課題状況で、タブレット版検査と紙ベース検査とでの再検査信頼性の評価を行い、両検査法での信頼性に差異があるかを確認する。

2. 方法

2.1 対象者

本研究の調査対象者は 49 名の大学生 (男性 33 名、女性 16 名) で、平均年齢は 19.47 ($SD = .50$) 歳で、まず Pap. D-CAT と Dig. D-CAT を行い、一定期日後に同様の手

続きで 2 回目を実施した。すべての対象者は視力や手指運動機能を始めとする認知能力や身体機能に問題がない大学生であった。幅広い年齢層での再検査を実施して検討することが最善であるが、高齢者で生じる可能性が高い体調の変化により影響を考慮しなくても良いことを考慮して、若年層を対象者とした。

2.2 検査材料と検査方法

Pap. D-CAT は、現在市販されている OCR での採点処理が可能な A 4 サイズ版の検査用紙を用いた。各列 45 の数字 (ゴシック体 18 point) が 14 行印刷されている検査用紙 3 種類 (数字の配列が異なり、各列に出現する 1 ~ 9 の数はランダムな順序で配列されている) の中から 2 種をランダムに選択して使用した。

検査者は対象者に筆記具を手渡し、「今から、私が指定した数字を 1 分間にいくつ消せるかを調べます。左上の数字から探し始め、指定された数字が見つかったら、できるだけ早く正確に、その数字を斜線で消してください。1 行目が終わったら、右の空欄に○印をつけ、次の行の左端に移ってください」と口頭で教示した。第 1 試行 (1 文字抹消課題) は「6」、第 2 試行 (3 文字抹消課題) は「8, 3, 7」の数字を抹消ターゲットとした。本研究での検討では、指標 TT と TO を用いることとした。FA について、そもそも、健常者ではごく僅かな値であり変量とみなすのは適当ではないため、これ以後の分析には加えない。実際、本研究では課題の理解不足やタブレット版の非意図的操作などによる外れ値を除外するため、スミルノフ・グラブス検定を行ったが、その際、FA は 1 でも外れ値として除外しなくてはならなくなったため、今回の分析からは除外した。

Dig. D-CAT は、サイズ 12.3 インチのタブレット画面に Pap. D-CAT と同じゴシック体 18 point 数字配列の画面が提示された。氏名、年齢などの対象者の属性を画面に記入したのち、「練習」画面が提示され、抹消手続きの説明が提示され、理解したと判断した場合は、画面上の「次へ」の表示をタッチして、検査画面へと移行した。ターゲット数字への反応は、タッチペンで行った。予備検査段階では、70 歳代の高齢者 10 名程度を対象に、指押し反応を採用したが、指先 (母指) の大きさ、押す力の入れ具合、などに困難さやばらつきが大きく、不適と判断してタッチペンでの反応を採用した。なお、タッチペンは Babble Com (P-001) 社製である。Dig. D-CAT は Windows ベースで起動し、(株) FIS が我々の求める仕様で作成したものを、数度の調整を経て作成した Version 4 をプロトタイプとして今回の検討に使用した。Dig. D-CAT は 60 秒条件、30 秒条件から選択できるが、本研究では Pap. D-CAT と同じく、60 秒条件を選択して実施した。

Dig. D-CAT での検査は個人別に机の上に寝かせてタブレットを置いて、画面に提示される指示に基づいて実施した。Pap. D-CAT と Dig. D-CAT の実施順序は対象者間でランダムとなるように配慮した。第 1 回目の検査と 2 回目の検査は 7 日間の時間間隔をおいて実施した。なお、2

回の検査は同じ場所で行った。

3. 結果

3.1 抹消数 (TT) に関する信頼性

信頼性係数を算出する前に、抹消数が 0 というデータが存在したため、Pap. D-CAT と Dig. D-CAT ごとにスミルノフ・グラブス検定を行い、1% 水準で異常値とみなせるデータを排除した。また、1 名の対象者は Pap. D-CAT の一部に欠損がみられたため、Dig. D-CAT のみ採用した。その結果、Pap. D-CAT と Dig. D-CAT の TT に関する分析の対象者数はともに 47 名であった。

表 1 は、Pap. D-CAT と Dig. D-CAT の TT 平均値と標準偏差を示している。

検査-再検査による信頼性係数は、Pap. D-CAT の 1 文字抹消課題では .78、3 文字抹消課題のそれは .83 であった。一方、Dig. D-CAT の 1 文字抹消課題における再検査信頼係数は .58、3 文字抹消課題のそれは .45 であった。

両検査の信頼性係数を比較すると、1 文字抹消課題における Pap. D-CAT と Dig. D-CAT の差は有意傾向で ($Z = 1.74, p = .08$)、Pap. D-CAT は Dig. D-CAT よりも再検査信頼性係数が高かった。3 文字抹消課題においても同様に Pap. D-CAT は Dig. D-CAT よりも再検査信頼性係数が高かった

($Z = 3.34, p < .01$)。再検査での類似度の理解を容易にするために、それぞれの検査での散布度を図 1 ~ 4 に示す。なお、第 1 試行は 1 文字抹消課題、第 2 試行は 3 文字抹消課題を示している。

3.2 見落とし数 (TO) に関する信頼性

TT と同様に、Pap. D-CAT と Dig. D-CAT ごとにスミルノフ・グラブス検定を行い、1% 水準で異常値とみなせるデータを排除した。その結果、Pap. D-CAT は 42 名、Dig. D-CAT は 45 名を分析の対象とした。表 2 は、Pap. D-CAT と Dig. D-CAT の TO に関する平均値と標準偏差を示している。

TO について、Pap. D-CAT の 1 文字抹消課題における再検査信頼係数は .15、3 文字抹消課題のそれは .47 であった。つぎに Dig. D-CAT の 1 文字抹消課題における再検査信頼係数は .26、3 文字抹消課題のそれは .32 であった。両検査の信頼性係数を比較した結果、1 文字抹消課題と 3 文字抹消課題のいずれにおいても有意な差は示されなかった。

4. 考察

本研究は 2 種類の D-CAT 検査の信頼性についての検討

表 1 : Pap. D-CAT と Dig. D-CAT の TT に関する平均値と標準偏差

	Pap. D-CAT		Dig. D-CAT	
	1 回目検査	2 回目再検査	1 回目検査	2 回目再検査
1 文字抹消課題	33.00 (5.06)	36.62 (5.31)	33.40 (5.77)	34.57 (5.62)
3 文字抹消課題	60.83 (9.56)	66.96 (12.02)	59.51 (9.91)	60.10 (11.30)

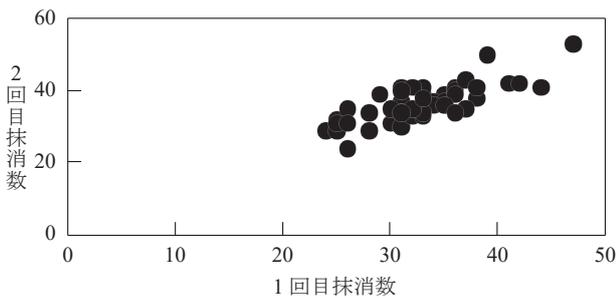


図 1 : Pap. D-CAT (第 1 試行)

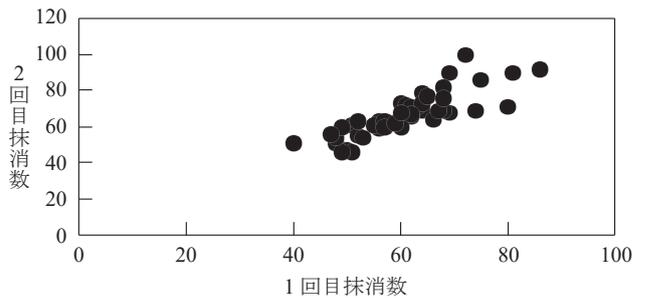


図 2 : Pap. D-CAT (第 2 試行)

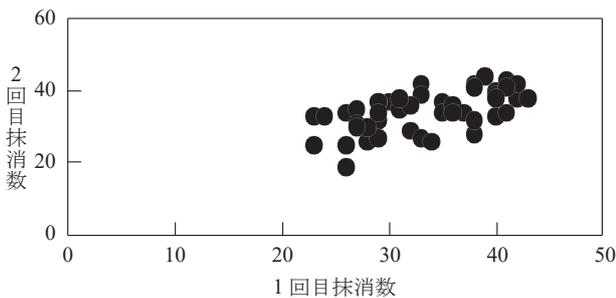


図 3 : Dig. D-CAT (第 1 試行)

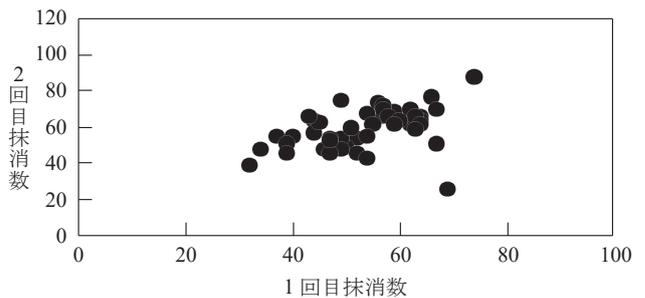


図 4 : Dig. D-CAT (第 2 試行)

表 2：Pap. D-CAT と Dig. D-CAT の TO に関する平均と標準偏差

	Pap. D-CAT		Dig. D-CAT	
	1 回目検査	2 回目再検査	1 回目検査	2 回目再検査
第 1 試行	0.43 (0.55)	1.29 (1.80)	1.42 (1.67)	1.82 (2.05)
第 2 試行	3.17 (2.72)	3.45 (3.07)	3.58 (3.84)	3.56 (2.66)

が目的であった。小塩 (2016) は本邦の査読誌論文を精査し、信頼性に関わる問題を整理している。信頼性に関連する要因には 2 回の検査の時間間隔と、信頼性計数値の大きさがあるとしている。まず、検査—再検査の時間間隔については、時間間隔の長短で再検査信頼性係数が大きく異なるという確認は得られなかったとしている。たとえば、異なる間隔で再検査信頼性を検討した安藤ら (1999) では、間隔が 3 週間よりも 17 週間の方が低下する傾向にあるものの、高本・服部 (2015) は、検査—再検査間隔を 1 週間から半年以上まで 7 段階にわけて再検査信頼性係数との関連を検討したが、期間が長くなるにつれて係数が低下するという明確な関連は見出していない。本研究の対象は質問紙検査ではないが、再検査までの時間間隔が 1 週間で短すぎるという懸念は取り上げるほどではないことになろう。高本・服部 (2015) は日本国内の心理尺度構成論文での信頼性係数報告を体系的にレビューし、再検査信頼性係数は検査—再検査の間隔によって多少異なるものの、おおむね .70 以上の値を示していることなどを報告している。小塩 (2016) はまた、望ましい信頼性レベルはどの辺りかについて評論し、その結果は一様でないとし、たとえば、.70 程度で十分で、尺度が、一時的な感情状態を反映するものであれば、低い方が望ましいとも言えるので、数値がどのような概念を反映するかの検討が重要であると述べている。

本研究での再検査相関係数は、分析結果、ターゲット数が 1 文字と 3 文字の両方において、Pap. D-CAT の信頼性係数は Dig. D-CAT のそれよりも有意に高いことが示された。その理由として明白なのは、検査の実施媒体である。紙媒体である Pap. D-CAT の検査—再検査による信頼性係数は、1 文字抹消課題で .78、3 文字抹消課題で .83 と十分な信頼性を許容できる範囲であるが、タブレット端末を用いた Dig. D-CAT に置き信頼性係数は 1 文字抹消課題が .58、3 文字抹消課題が .45 と若干低いと言わざるを得ない値であった。なお、310 名の大学生・短大生を対象者とした Hatta et al. (2014) による紙媒体の再検査信頼性の検討では、1 文字抹消課題での再検査信頼性係数は .81 で、3 文字抹消課題の場合は .75 あった。再検査までの期間は 2 週間である。両研究結果の差異は検査期間の長さが違うことがもたらした可能性はあるが、ほぼ無視できるレベルの違いで、Pap. D-CAT は十分に高い信頼性を有しているといえよう。

Dig. D-CAT と Pap. D-CAT では同様の作業を行うことから課題そのものに本質的な差はない。Pap. D-CAT では高い信頼性が示されていることから、これらの信頼性係数

に関する差は、媒体による作業効率や正確性の問題であると考えられる。具体的には、タッチパネル上でターゲットとなる数字にペン先を触れさせるという作業工程と、紙媒体のターゲットとなる数字にペンで斜線をひくという作業工程の違いや、タッチパネルではタッチペンで触れても反応が悪く繰り返しタップしたり、少しのずれで意図した数字ではない周辺の数字にチェックがついてしまったりするという正確性についての問題が想定できる。

正確性への差異をもたらす理由として、タブレット画面の数字をタッチペンで押す動作とペンで抹消線を引く動作との筋運動系での微妙な違いが考えられる。タッチペンで画面を押す動作は、筆圧の加減を常にコントロールする要素が含まれる。ペン先を「強く押しすぎない」ように筆圧を調整 (抑制) する要素は、手先の筋運動系へのかなりの負荷となる。そして、この筆圧調整コントロールは検査開始時に初めて経験するものであり、経験値の高い、つまり手続き学習の自動化による影響が大きい紙に印刷された数字に斜線を書く動作に比べると、負荷が大きいことが起因する可能性がある。同じ作業工程で同様の課題を繰り返し行っている以上、Dig. D-CAT と Pap. D-CAT の信頼性の違いは、反応過程での正確性の問題につながる。つまり、Dig. D-CAT は Pap. D-CAT よりも反応時の均一性が低いため、個人内での調整がうまくいった場合とそうでない場合に差が生じやすく、信頼性係数がやや低かったと考えられる。実際、Dig. D-CAT と Pap. D-CAT で TT の値に大差はないことから、課題自体の問題ではなく、媒体におけるターゲットを抹消する際の反応の均一性の問題であると推察できる。

この問題を解決する手段としては、2 つの方法が考えられる。1 つ目は均一性、正確性を補うための装置の改善である。しかし、これは電子工学技術上の問題であり、それを改善することは心理学の領域外であると言わざるを得ない。二つ目は、こうした媒体による正確性を補正することである。この方法は先に挙げた媒体の違いにおける作業工程の違いを補う上でも有効である。今後の課題は、Dig. D-CAT の信頼性を Pap. D-CAT のそれにより近づけるべく補正の方法を見出すことであろう。

本研究では、TT 以外にも Dig. D-CAT と Pap. D-CAT の信頼性係数を比較する指標として TO についても検討した。分析の結果、ターゲット数が 1 文字と 3 文字のいずれにおいても Pap. D-CAT と Dig. D-CAT で有意な差は示されなかった。このことから、見落としというミス反応について、媒体による違いを考える必要はないのかもしれない。TT と TO で媒体の違いがもたらす結果が異なる理

由は TT に関する考察から説明可能である。つまり、TT では、タブレット画面上の数字をタップする際の均一性・正確性が差を生じさせる原因であると想定されるが、TO の場合、抹消すべき数字を見落とすことを意味するため、ペンでタップすべき画面上の数字をタップしないことで生じるミス反応の数である。それゆえ、媒体による操作の均一性・正確性という問題は生じないため、媒体による差は示されなかったと解釈できる。

媒体の差は合理的に説明可能であるとして、本研究の結果で注目すべき問題点は、TO の信頼性係数が Pap. D-CAT の 1 文字抹消課題で .15、3 文字抹消課題で .47、Dig. D-CAT の 1 文字抹消課題で .26、3 文字抹消課題で .32 と低いことである。その背景には D-CAT という課題の性質があるかもしれない。

課題の中身は乱数の文字列からターゲット数字を見つけ抹消することで、健常成人にとって難易度は極めて低い。そのために、抹消すべき数字を見逃す度数は低いために 1 回目の検査では生じず、2 回目で生じたというような場合が多く、つまり同一人が同様のミスを行うことが稀なために相関係数は低くなったと推定できる。頭部外傷患者でのように、脳の機能障害で見落とす頻度が多くなる場合には 1 回目と 2 回目の検査事態で類似の見逃し反応が増えて、一定の相関係数がもたらされる場合とは異なる結果に至った可能性が考えられる。つまり、本研究の対象者が健常な大学生であり頭部外傷患者や注意機能低下が顕著になり始める高齢者群ではなかったために、見落とすというミス反応はごくわずかで、変量とみなしうるかに疑問がある水準である。ただ、D-CAT 開発は、頭部外傷患者の注意機能を評価するのが根本的な目的であり（伊藤・吉崎・八田, 2001; Ito & Hatta, 2003）、注意機能の低下が顕在化する対象、例えば高齢者、頭部外傷者での評価が目的であるために、健常者での TO 値は変量の要素を満たさないことは自明と言えるかも知れない。もっとも、タブレットをはじめとする情報端末の操作については、加齢による影響を受けやすく、端末利用時における高齢者特有のエラー操作であったり作業内容の認知的負荷が高いとエラーが多くなることは報告されており、これらの要因を加味してさらなる検討を加える必要がある（緒方・上田・須藤・熊田・伊福部, 2012）。本研究では、Dig. D-CAT の検査 - 再検査の検討はそれほど大きな母集団で行われていないこと、対象者の検査場面への動機づけのレベルに問題がなかったか、対象者の背景にばらつきが乏しいことなどを考慮してさらに検討する必要がある、直ちに Dig. D-CAT の再検査信頼性についてその評価を結論づけることは時期尚早であろう。

謝辞

本研究は科学研究費助成事業、挑戦的萌芽研究 17K18716（代表者：八田武志）及び基盤研究（C）20K03423（代表者：八田武俊）の助成を得て実施した。また、本研究は関西福祉科学大学研究倫理審査委員会により承認（承認番号 22-21）を経て実施した。資料収集において、梅本篤樹、新宮春輝、

竹村龍、田中大夢、玉置直也、寺田侑生、野々山優吾、船越颯真、山本葉大の諸氏に協力を得たことを記し、謝意を表します。

引用文献

- 安藤明人・曾我祥子・山崎勝之・島井哲志・嶋田洋徳・宇津木成介・大芦治・坂井明子 (1999). 日本語版 Buss-Perry 攻撃性質問紙 (BAQ) の作成と妥当性, 信頼性の検討. *心理学研究*, 70, 384-392.
- 八田武志 (2007). 神経心理学研究における測定と評価について. *神経心理学研究*, 23, 2-7.
- Hatta, T., Yoshizaki, K, Ito, Y., Mase, M, & Kabasawa, H. (2014). Reliability and validity of the Digit Cancellation Test: A brief screen of attention. *Psychologia*, 55, 246-256.
- 八田武志・伊藤保弘・吉崎一人 (2014). D-CAT—注意機能スクリーニングテスト使用手引き—. FIS.
- Itoh, Y. & Hatta, T. (2003). Re-analysis of the performance of traumatic brain injury patients on the available neuropsychological tests. *Psychologia*, 46, 199-209.
- 伊藤保弘・吉崎一人・八田武志 (2001). 頭部外傷 (TBI) 患者の情報処理速度の低下および注意障害—簡便評価法 (D-CAT) による検討—. *情報文化研究*, 14, 43-57.
- 緒方啓史・上田一貴・須藤智・熊田孝恒・伊福部達 (2012). 加齢による認知機能変化が高齢者の ICT 機器を用いた就労への意欲に及ぼす影響. *情報処理学会論文誌*, 53, 1696-1710.
- 小塩真司 (2016). 心理尺度構成における再検査信頼性係数の評価. *心理学評論*, 59, 68-83.
- Sohlberg, M. & Mateer, C. A. (1989). *Introduction to cognitive rehabilitation: Theory and practice*. New York: Guilford Press.
- 高本真寛・服部環 (2015). 国内の心理尺度作成論文における信頼性係数の利用動向. *心理学評論*, 58, 220-235.

(受稿：2022 年 10 月 9 日 受理：2022 年 11 月 25 日)