

内田クレペリン検査の鉄道事故および ヒューマンエラーに対する識別性の評価¹⁾

楠 神 健*

Discrimination Power of the Uchida-Kraepelin Psychodiagnostic Test in Relation to Railway Accidents and Human Error

Ken KUSUKAMI

The Uchida-Kraepelin psychodiagnostic test is a personality test assessing testee's ability to work and the characteristics of his/her work process through his/her performance of successive one-digit addition. In Japan, railway drivers are required to pass this test periodically to prove their safety performance. In this research, discrimination power of this test in relation to accidents and human error is assessed by analyzing test scores and accident data of 6018 railway drivers. Discrimination power of the test in relation to accidents is suggested in drivers with less than three years of experience, but not so with the other drivers. The cause is considered to be that the ability of distribution and sustainability of attention measured by the test directly influences driving performance of less experienced drivers. The discrimination power of the test in relation to human error types is not proven in this analysis.

key words: Uchida-Kraepelin psychodiagnostic test, aptitude test, railway driver, accident, human error

問 題

内田クレペリン検査は人格検査あるいは作業検査の一種とされ、1分単位の連続加算作業を休憩を挟み30回繰り返す、その間の作業量の変動および計算誤りから、作業ぶりや作業能力を評価する検査である(日本・精神技術研究所, 1975)。日本では広く活用され、企業では採用や安全管理等に、病院では診断や治療効果測定等に、学校関係では入学試験や教員の採用試験等に用いられている(瀧本, 2013)。

鉄道では、日本国有鉄道において、速度違反や信

号違反等の多発から、事故の未然防止のために1949年に運転適性検査体系が制定されて以来(大庭, 1981)、運転適性検査の1項目として使用されてきた。また、1987年の日本国有鉄道の分割民営化以降も、省令等の定めにより、私鉄も含め鉄道の運転士の資格検査として現在に至るまで用いられている(喜岡, 2000)。なお、検査は、採用時(運転適性検査の枠組みではなく採用時の検査として実施)、運転士登用時(臨時検査)および登用後は3年に1回以上(定期検査)受けることになっている。

一方、1949年以降、運転環境や運転条件は大幅

¹⁾ 本研究については、野口裕之教授(名古屋大学大学院教育発達科学研究科)、長塚康弘名誉教授(新潟大学)、芳賀繁教授(立教大学文学部)、渡邊忠教授(文教大学人間科学部)、吉田信彌教授(東北学院大学教養学部)から研究推進上の様々なアドバイスをいただきました(ご所属等は本研究実施時のものです)。各先生に深く感謝の意を表します。

* 東日本旅客鉄道株式会社 JR 東日本研究開発センター
Research and Development Center of JR East Group, East Japan Railway Company
2-479 Nisshin-cho, Kita-ku, Saitama 331-8513, Japan
e-mail: kusukami@jreast.co.jp

に変化している。たとえば、運行される主な車種は蒸気機関車から電車に変わり、それに伴い、運転士の作業は、機関等を制御する運動機能主体の作業から信号や外乱等の監視を行う精神的な作業が主体となり、質的な変化が発生している。また、この間、ATS (Automatic Train Stop) や ATC (Automatic Train Control) 等の運転士のヒューマンエラー（以下、単にエラーという場合、ヒューマンエラーを意味する）を事故に直結させない信号保安装置が整備され、鉄道の安全システムの仕組みも変化している（喜岡，2000；楠神，2013）。

そのため本研究では、運転士の運転環境等の変化を経ても、内田クレペリン検査が事故やエラーと関連を持っているかどうかの確認を行う。

ここで、まず内田クレペリン検査の結果の評価方法について簡単に説明する。本検査は、作業量および作業量の変動と計算誤りから評価を行うが、運転士登用時以降は作業量および誤答数で検査不合格となる者がほばいないため、本研究では作業量の変動のみを評価の対象とする。これは、1分単位の加算作業の最終到達点を線で結び、その曲線の形状（プロフィールという）から評価を行う（日本・精神技術研究所，1975）。

プロフィールの判定には、大きく直観判定と数量判定の2つの方法がある。直観判定とは、1分単位の作業量に応じて作業量の段階をA～Dの4つに区分（鉄道総合技術研究所では4段階だが、日本・精神技術研究所では前記A段階をさらに2つの段階に区分している。分割民営化後のJR各社では前者の区分で実施しているため、本論文ではこの区分に沿って記述する）したうえで、曲線の形状の判定を検査員が行う方法である。判定結果は、曲線類型という離散的な判定区分で示され、たとえば作業量がA段階の場合、定型的な曲線（基準となる曲線）からのずれの程度および非定型特徴（激しい動揺、後期作業量の下落、後期初頭の出不足などいくつかの種類に分けられる。日本・精神技術研究所，1975）の程度と数に応じて6段階に区分される。これが内田クレペリン検査における一般的な判定法である。

一方、ソフトウェアによるPF値等を用いた数量判定も可能になってきており、JR各社では、数量判定結果を踏まえて、検査員が最終的に直観判定により曲線類型を判定する方法に移行している。PF

値とは、Profile Fluctuation（曲線変動）の略で、作業各行における各個人の実際の作業量と期待作業量のずれの2乗量を標準化したものの総和であり、期待作業量は、当該個人の各分作業量の平均から回帰予測される（柏木，1975）。PF値は、プロフィールが定型的な曲線に近いほど数値が小さく、逆に非定型特徴の数が多く、その程度が著しいほど数値が大きくなる。このようにPF値は曲線類型とは異なり、定型的な曲線からのずれの程度が連続量として測定可能である点が特徴である。

検査成績と事故等との関連の分析（分析1）

まず本研究では、運転士の内田クレペリン検査の成績と事故等との関連について分析する。これまで、運転士に限った分析はあまり行われていないが、導入当初の分析では、相馬（1949）、鶴田（1950）とも両者に関連があるとの結果であった。なお、鶴田（1950）については原典が入手できなかったため、三隅・白樫・安藤・黒川（1961）のレビューの中で確認した。後者は曲線類型の分布の違いに基づく分析、前者は事故頻発性指数という独自の指標による分析である。その後も、柏木（1964）、鉄道労働科学研究所名古屋駐在（1970）、日本・精神技術研究所（1975）、藪原（1986）、藪原（1990）と断片的に分析が行われている。柏木はPF値、鉄道労働科学研究所名古屋駐在、日本・精神技術研究所では曲線類型の分布、藪原の2つの研究では、事故指数という独自の指標で分析していた。前の3つの研究については、事故群は無事故群よりも成績が悪い傾向にあった（ただし、鉄道労働科学研究所は有意差なし、他は検定はされていない）。一方、藪原の2研究については、成績と事故指数との対応が十分見られなかった。それ以降は、運転士に関する検査成績と事故との関連の分析はみられない。したがって、本研究では、あらためて運転士について両者の関連の分析を行う。

また、関連の分析方法としては、事故者と無事故者の成績の差をみることにした（ここではこの差を検査の識別性と定義する）。内田クレペリン検査の事故に対する基準関連妥当性を評価するためには、検査成績の変動に伴う事故者の割合など、検査の予測力をみる必要があるが、本論文では以下の理由により、識別性の評価を行うことにした。第一に、前記のように運転士は採用時や運転士登用時に内田ク

レペリン検査に合格しているため、検査成績と事故との関連は、検査合格者に対してしか評価できない。したがって、検査指標の本来の予測力については評価ができない。次に、運転適性検査の目的は、事故者と無事故者の識別であるため、識別性が最もそれを直接的に測定する方法である。

一方、これまでの研究では、評価の対象である“事故”の指標についても、影響が「列車の運休や遅れ」のみの輸送障害（定義の詳細は鉄道事故等報告規則を参照）が主体であり、列車の脱線・衝突等、旅客の死傷等につながる鉄道運転事故（同様に鉄道事故等報告規則を参照。以下、単に事故という）に対する妥当性の検証は行われてこなかった。これは運転士のヒューマンエラーに起因する事故がきわめて少なく、エラーの結果が列車の遅れのみの輸送障害になることが大部分であることに起因する。

したがって、本研究ではこの点についても、評価対象を輸送障害のみではなく、事故に発展する可能性のあったリスクの高い事象（インシデント、詳細後述）にも着目する。両者を区別した理由は、以下の通りである。これまでリスクの高い事象（停止信号を行き過ぎる事象など）には優先的にシステム対策（信号保安装置など）が打たれてきたため、これらの事象が発生する際には何らかの異常時エラーが関与することが多くなっている。異常時エラーには情動が影響しやすいとされるため（細田・井上, 2000）、異常時エラーの発生には平常時エラーとは違う心理特性が関与する可能性があると考えたためである。

検査成績とエラーとの関連の分析（分析2）

内田クレペリン検査が、事故等との関連がある場合、その原因を探るのが次の目的である。本検査のプロフィールが意味する内容については、検査結果と(a)特定のヒューマンエラーを起こしやすい課題の結果との関連（井上・重森・喜岡・赤塚・宮地, 2006）、(b)注意の特性を測る検査の結果との関連（喜岡・宮地・山内, 2003; 喜岡・鈴木・宮地, 2004）、あるいは(c)既存の性格検査の結果との関連（柏木, 1995; 瀧本, 2000; 鈴木・喜岡・宮地, 2004）からエラー、注意、あるいは性格との関連が一定程度分析されている。

それらを含め、内田クレペリン検査によって測定

されている特性（以下、測定特性という）を検討するため、ここでは分析1で用いたインシデントについて、それを引き起こしたヒューマンエラーの内容を分析し、エラーのタイプと内田クレペリン検査の結果との関連を分析する。ヒューマンエラーによる事故を防ぐために検査が行われているので、発生しているエラーの内容と検査結果の関連を分析することがまず重要だからである。

方 法

検査の事故等に対する識別性の評価方法（分析1）

評価に用いる検査指標 識別性の評価に用いる検査指標はPF値とした。その理由は、曲線類型がたとえば作業量A段階でいえば6段階のカテゴリ変数であるのに対し、PF値は連続量であり、曲線類型に比べ情報量が多いためである。

検査によって予測する事象 東日本旅客鉄道株式会社（以下、JR東日本という）では、輸送障害の中でも、事故に発展する可能性のあったリスクの高い事象（国の分類上は輸送障害。以下では、インシデントと呼ぶことにする）を別に分類している。本研究では、件数が極めて限定される事故のデータに加え、インシデントを事故に発展する可能性のある事象として用い、輸送障害経験者に加え、インシデント経験者についても検査成績との関連をみる。

分析データ 識別性の評価に用いるデータについて説明する。

1. 分析対象運転士および検査成績データ: 分析の対象とした運転士は、JR東日本管内において、2001年4月1日現在で在来線の本線の運転業務に携わっている社員（以下、単に運転士という）であり、有効サンプル数はこの段階で7,363名であった。次に、これらサンプルについて上述した曲線類型の作業量区分をみたところ、97%以上がA段階であったため、識別性の評価対象もA段階の運転士のみとした。また、数量判定によりPF値が得られているサンプルのみが評価対象になるため、分析に用いた運転士は6,018人になり、また数量判定の導入期であるため、PF値のデータは1997~2000年度の定期検査データになった。なお、上記の取扱いにより、事故発生時と検査時に最大14年の差が生じる。しかし、適性検査で測定する心的特性は比較的に変わりにくい先天的な素質であり（鉄道労働科学研

究所心理適性管理室, 1987), また, 定期検査を実施する意義は, 登用後の著しい機能低下を来した少数例を発見するためとされている(鉄道労働科学研究所心理適性管理室, 1980)。したがって, 本研究ではこの前提に基づき上記方法により分析を実施した。

分析対象運転士の2001年4月1日現在の年齢平均は39.6, 標準偏差は11.0, 運転士経験日数の平均は3513, 標準偏差は1847, 性別は全員男性であった。なお, 前述のように, 鉄道における適性検査は資格検査であるため, 検査不合格者は運転士にならないこと, つまり分析は検査合格者に対してのみ行われたことを追記する。

2. 分析対象とする輸送障害等の事象: 1987年4月1日から2001年3月31日の14年間に発生した運転士のヒューマンエラーが関与する事故・輸送障害のうち, 上記の分析対象運転士6,018人が当事者でかつ検査成績との対応付けが行えたものを分析対象にした。該当運転士は事故・輸送障害経験者が2,218人(3,523件)。なお, そのうち事故は2件のみで残りは輸送障害。以下, 単に輸送障害経験者という), そのうち, 事故・インシデント経験者は247人(251件。事故は上述の2件。以下, 単にインシデント経験者という)であった。その結果, インシデント経験者以外の輸送障害経験者(以下, 低リスク輸送障害経験者という)は1,971人, 事故およびインシデントを含む輸送障害を全く起こしていない者(以下, 完全無事故者という)は3,800人になった。この3カテゴリによる分類を事故ランクによる分類と呼ぶ。

分析方法 事故ランク別の分析については, 完全無事故者, 低リスク輸送障害経験者, インシデント経験者の3水準による1要因分散分析を実施し, 識別性を評価する。また, 多重比較の結果, 今回分離した低リスク輸送障害経験者とインシデント経験者との間にPF値の差が認められない場合には, 両カテゴリを統合した輸送障害経験者と完全無事故者との間で平均値の差の検定を行う。

また後述するように, サンプル全体だと3つのカテゴリ別の運転士経験日数の平均値に差がみられたため(カテゴリ間で1,000日程度の差がある), JR発足時にすでに運転士で2000年度まで14年間運転士であった群(以下, 現職14年という。ただし国

鉄時代から継続して運転士だったと考えられるため, 実質的には現職14年以上になる)の分析を行った。一方, 現職14年未満のサンプルについても同様の理由から, 群のサンプル数を保ちながら経験日数の平均値の差を減らすため, 現職が3年未満, 3年以上6年未満, 6年以上10年未満, 10年以上14年未満の4群に区分したうえで, 同様の分析を行った。

検査のエラーに対する識別性の評価方法(分析2)

ヒューマンエラーの分類方法 次に, 内田クレペリン検査が上記事象に識別性を持つ場合の原因を探るため, 内田クレペリン検査がどのようなタイプのヒューマンエラーに識別性があるのかをみる。エラーのタイプ分類としては, Rasmussen(1983)のヒューマンパフォーマンスの概念を取り入れたReason(1990)の分類を用いる。具体的には, スキルベースのスリップ(または単にスリップという), スキルベースのラプス(または単にラプスという), ルールベースのミステイク, ナレッジベースのミステイクの4分類である。

この分類を用いた理由は以下のとおりである。ここでエラータイプ別の識別性をみる目的は, その分析結果を通して, 内田クレペリン検査が事故等に識別性を持つ原因あるいは測定している心理特性を探るためである。Reasonのエラー分類は, 作業の性質や作業者の慣れの程度を反映した情報処理プロセスの違いおよびそれを踏まえたエラー発生メカニズムの違いによりエラーを分類している。したがって, 当該検査の測定している心理特性等を分析するのに適していると考えたためである。また, 英国の鉄道事故等の分析に用いられるなど, 鉄道のエラー分析についての実績もある(Gibson, Smith, Lowe, Mills, Morse, & Carpenter, 2013)。

インシデント発生に関連するエラーの分析 分析対象としたエラーは, インシデントに対する識別性の評価に用いたインシデント251件(247人)であり, その中で発生した運転士のヒューマンエラーをすべて抽出したうえで(抽出エラーは全部で455), それらを前述の4つのエラータイプに分類した。なお, インシデントで発生したエラーのみを分析対象にした理由は以下の通りである。前述のように, 事故につながりうるインシデントを惹起するエラーは, 低リスク輸送障害のそれよりも異常時エラーの

割合が高いなど、質的に異なると考えられる。したがって、鉄道事業の安全にとって最も重要な「事故の防止」に繋がるインシデントを対象に分析することにした。エラーの抽出と分類は、JR 東日本の事故報告書に基づき実施した。その方法は、運転作業およびそこで発生する事故・エラーに精通し、かつ心理学の素養のある3名がまず個別にエラーの抽出・分類を行ったうえで、1件ごとにエラー抽出および分類について、再度、事故報告書の内容を吟味しながら合議のうえで決定した。

分類の結果、スキルベースのスリップと判断されたのが165エラー(130人)で全エラーの36%を占め、最も多かった。信号保安装置の進展により、うっかりミスが事故に繋がる可能性は大幅に減少しているが、池田・大嶽(1974)、鉄道労働科学研究所労働心理研究室・人間工学研究室・心理適性管理室(1982)、米山・宇賀神(1986)などと同様、インシデントに関連するエラーで最も多いのは依然としてスリップであった。一方、スキルベースのラプスが57エラー(50人)、ルールベースのミステイクが62エラー(50人)で、それぞれ全エラーの13%および14%を占めた。なお、ナレッジベースのミステイクと判断されたのは8エラー(6人)で全エラーの2%と少なく、識別性について安定した分析が困難であるため、今回の分析対象から外した。したがって、以下でミステイクという場合、ルールベースのミステイクを意味する。上記データにおいて、4つのエラータイプの件数の合計が全インシデント件数を上回るのは1つのインシデントに複数のエラーが発生するケースがあるからである。また、4エラータイプを経験した人の合計が全インシデント経験者を下回るのは、事故報告書の記述からエラータイプが判別できなかったエラー、4つのエラータイプとは異なると判断されたエラーがあったからである。

検査成績とエラータイプとの関連の分析方法 3つのエラータイプに対する識別性の評価方法は、前記の事故ランク別の分析に準じるが、今回はインシデントの中のエラーしか分析していないため、検査のエラー識別性をより正確にみるため、以下の3水準で比較した。すなわち、インシデントの中で当該エラータイプを経験した者(それぞれスリップ経験者、ラプス経験者、ミステイク経験者という)、イ

ンシデントでのエラー経験はあるが当該エラータイプは未経験の者(スリップ以外のエラー経験者、ラプス以外のエラー経験者、ミステイク以外のエラー経験者という)、事故・輸送障害を通してエラー経験のない者(以下、完全無エラー者といい、完全無事故者と一致する)である。なお、完全無エラー者と3つのエラータイプ経験者の4群で分散分析を行わなかった理由は、複数のエラータイプを重複して経験している運転士がいるからである。また、エラー経験はあるが当該エラータイプは未経験の者を水準に加えたのは、当該エラー経験者の成績を完全無事故者に加え、その他のエラー経験者とも比較するためである。また、サンプルの群の設定方法は事故ランク別分析と同一である。

結 果

事故ランク別の識別性の評価(分析1)

Table 1に事故ランク別の分析結果を全体および5つの群別に記す。表には、完全無事故者、輸送障害経験者、および後者を2つに分類した低リスク輸送障害経験者とインシデント経験者について、サンプル数、PF値の平均値と標準偏差、および参考として2001年4月1日現在の年齢および運転士経験日数の平均・標準偏差を記載した。PF値の平均はいずれも9前後から12前後で、これは定型でも非定型でもない領域であり、また、6段階の評定の中では上位から4段階に相当する成績である。

全体および5つの群別に一要因の分散分析およびTukeyの多重比較を行ったところ、全体では、低リスク輸送障害経験者およびインシデント経験者が完全無事故者よりもPF値が有意に高く($F(2, 6015) = 14.18, p < .01, \eta^2 = .005$)、現職3年未満では、低リスク輸送障害経験者が完全無事故者よりも有意にPF値が高かった(なお、分散分析の結果は、 $F(2, 1058) = 2.96, p = .052, \eta^2 = .006$ で、有意ではなかったが、多重比較で有意差が認められた)。それ以外の群では有意差が認められなかった。

一方、全体および5つの群別の多重比較結果から、すべての分析で低リスク輸送障害経験者とインシデント経験者のPF値に差が認められなかった。今回、リスクの程度により、低リスク輸送障害とインシデントを区分して分析したが、両者のPF値に差がないことがわかった(詳細は後に考察する)。

Table 1 事故ランク別，現職経験年数別の PF 値等の平均および標準偏差

		n	PF 値		年齢		経験日数	
			M	SD	M	SD	M	SD
全体	完全無事故者	3800	10.35	4.65	37.7	11.1	3127	1957
	輸送障害経験者	2218	11.02	4.90	42.9	10.0	4174	1508
	低リスク輸送障害経験者	1971	11.01	4.95	42.9	9.9	4169	1516
	インシデント経験者	247	11.12	4.55	42.9	10.3	4210	1450
現職 14年	全体 N=3123							
	完全無事故者	1622	11.25	4.85	46.8	6.3	5114	0
	輸送障害経験者	1501	11.29	4.89	47.8	6.3	5114	0
	低リスク輸送障害経験者 インシデント経験者	1330 171	11.27 11.43	4.93 4.60	47.7 48.1	6.3 6.5	5114 5114	0 0
現職 14年 未満	3年未満 N=1061							
	完全無事故者	926	8.92	4.02	25.4	4.7	558	308
	輸送障害経験者	135	9.83	4.53	25.6	4.5	738	237
	低リスク輸送障害経験者 インシデント経験者	123 12	9.84 9.77	4.60 3.87	25.6 25.8	4.5 5.2	725 865	242 140
3年以上 6年未満 N=985	完全無事故者	694	9.72	4.56	30.3	7.1	1679	302
	輸送障害経験者	291	9.91	4.47	29.5	6.3	1751	308
	低リスク輸送障害経験者	257	9.85	4.47	29.7	6.5	1740	308
	インシデント経験者	34	10.36	4.57	28.2	4.8	1833	301
6年以上 10年未満 N=561	完全無事故者	361	11.10	4.45	40.0	8.0	2880	409
	輸送障害経験者	200	11.38	5.70	38.2	7.1	2903	408
	低リスク輸送障害経験者	178	11.53	5.85	38.6	7.1	2916	413
	インシデント経験者	22	10.16	4.12	35.4	7.4	2794	358
10年以上 14年未満 N=288	完全無事故者	197	10.47	4.46	42.7	5.4	4392	432
	輸送障害経験者	91	11.21	4.45	41.8	5.4	4306	477
	低リスク輸送障害経験者	83	11.11	4.37	42.0	5.6	4345	477
	インシデント経験者	8	12.19	5.45	38.8	1.4	3910	257

注1: 現職14年の運転士の経験日数の分散が0になっているのは，全運転士がJR発足時に運転士であり，経験日数を一律にJR発足日から起算したためである。

したがって次に，両者を統合して，輸送障害経験者と完全無事故者の PF 値について *t* 検定を行った。

まず，分析対象者全体の PF 値をみると，完全無事故者 (10.35) よりも輸送障害経験者 (11.02) の方が有意に PF 値が高かった ($t(6016)=5.32, p<.01$, Cohen の $d=0.14$)。一方，完全無事故者よりも輸送障害経験者の方が年齢・運転士経験日数とも有意に高かった (年齢については， $t(6016)=18.88, p<.01$, Cohen の $d=0.49$ ，経験年数については， $t(6016)=23.22, p<.01$, Cohen の $d=0.58$)。2001 年度初の新入運転士を対象としているため，経験の長い運転士ほど，輸送障害に遭遇していることがわかる。

次に，もっとも条件が統制されている現職14年の運転士 (ただし，当然ながらサンプル自体は全運転士に比べ年齢の高い方に偏っている) の *t* 検定結果をみると，完全無事故者 (11.25) と輸送障害経験者 (11.29) との差がほとんどなくなった ($t(3121)=$

0.21, $p=.94$)。また，これらの運転士は 37~58 歳まで分布するので，それらを 35~39 歳のように 5 歳刻みでグループ化し，グループ別に *t* 検定を行っても有意差は認められなかった。

一方，現職14年未満をみると，有意差が認められたのは，現職3年未満のみで，完全無事故者 (8.92) よりも輸送障害経験者 (9.83) の方が有意に PF 値が高かった ($t(1059)=2.44, p<.05$, Cohen の $d=0.22$)。

以上，全サンプルで分析すると，完全無事故者と輸送障害経験者の間に，PF 値の他，経験日数や年齢でも差がみられたことから，現職経験年数等の層別化により経験量を統制しながら分析を行った。その結果，内田クレペリン検査は，全体に加え，現職3年未満の群に対して完全無事故者と輸送障害経験者の間に PF 値に差がみられた。一方で，それ以外の群では有意な差がみられず，特に最も条件が統制

されている現職14年のサンプルではPF値に差がみられなかった。このことから、一定の運転経験を積むと、内田クレペリン検査の測定特性については、事故との関連が薄れる可能性がある。

一方、有意差が確認された経験年数の浅い現職3年未満であっても、各水準のPF値の標準偏差を考慮すると、平均値の差はそれほど大きくなく、効果量(Cohenの*d*)も0.22でCohenの基準でいえば「小さい」効果量に相当する。したがって、両者に成績の差はあるが、あらかじめ予測するという視点に立てばその力が強いとまではいえなかった。これらの点については、考察の中で再度議論する。

エラータイプ別の識別性の評価(分析2)

Table 2に、3つのエラータイプに関する各水準別のサンプル数、PF値の平均値と標準偏差、2001年4月1日現在の年齢および運転士経験日数の平均・標準偏差を示す。群別の分析は、事故ランク別と同様、全体および5群で実施したが、現職14年未満のサンプルを細分化した4群は、エラー経験者のサンプル数が一桁になり、安定的な結果が得られなかったため、Tableには記載しなかった。

まず、分析対象者全体のPF値をみると、分散分析でスリップおよびミステイクで有意差が認められた(両者とも $F(2, 4044)=3.18, p<.05, \eta^2=.002$)。なお、Tukeyの多重比較では、両者ともすべての組合

Table 2 エラータイプ別、現職経験年数別のPF値等の平均および標準偏差

	<i>n</i>	PF 値		年齢		経験日数			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
全体	スリップ	完全無エラー者	3800	10.35	4.65	37.7	11.1	3127	1957
		スリップ経験者	130	11.08	4.85	44.0	10.1	4328	1380
		スリップ以外のエラー経験者	117	11.15	4.22	41.6	10.5	4079	1519
	ラプス	完全無エラー者	3800	10.35	4.65	37.7	11.1	3127	1957
		ラプス経験者	50	10.36	4.37	42.5	9.9	4254	1366
		ラプス以外のエラー経験者	197	11.31	4.59	43.0	10.5	4199	1474
	ミステイク	完全無エラー者	3800	10.35	4.65	37.7	11.1	3127	1957
		ミステイク経験者	50	11.19	3.74	41.9	11.3	4020	1616
		ミステイク以外のエラー経験者	197	11.10	4.75	43.1	10.1	4259	1405
現職 14年	スリップ	完全無エラー者	1622	11.25	4.85	46.8	6.3	5114	0
		スリップ経験者	94	11.39	4.86	48.6	6.4	5114	0
		スリップ以外のエラー経験者	77	11.49	4.29	47.6	6.5	5114	0
	ラプス	完全無エラー者	1622	11.25	4.85	46.8	6.3	5114	0
		ラプス経験者	34	10.77	4.57	47.7	6.3	5114	0
		ラプス以外のエラー経験者	137	11.60	4.61	48.2	6.5	5114	0
	ミステイク	完全無エラー者	1622	11.25	4.85	46.8	6.3	5114	0
		ミステイク経験者	32	11.74	4.13	48.7	6.7	5114	0
		ミステイク以外のエラー経験者	139	11.36	4.71	48.0	6.4	5114	0
現職 14年 未満	スリップ	完全無エラー者	2178	9.68	4.38	30.9	8.8	1647	1246
		スリップ経験者	36	10.29	4.81	31.9	7.7	2277	1016
		スリップ以外のエラー経験者	40	10.50	4.05	30.2	6.5	2088	826
	ラプス	完全無エラー者	2178	9.68	4.38	30.9	8.8	1647	1246
		ラプス経験者	16	9.48	3.90	31.4	6.8	2426	923
		ラプス以外のエラー経験者	60	10.65	4.52	30.9	7.2	2111	915
	ミステイク	完全無エラー者	2178	9.68	4.38	30.9	8.8	1647	1246
		ミステイク経験者	50	10.19	2.76	29.8	6.6	2075	1128
		ミステイク以外のエラー経験者	199	10.47	4.81	31.4	7.2	2209	854

注1: 現職14年の運転士の経験日数の分散が0になっているのは、全運転士がJR発足時に運転士であり経験日数を一律にJR発足日から起算したためである。

せで有意差なし。また、ラプスも分散分析上は有意差が認められたが、評価の中心になる完全無エラー者とラプス経験者の間に差はなかった)。一方、条件の最も統制されている現職14年の群でみると、事故ランクと同様、3エラータイプとも有意差が認められなかった(ラプスについてはむしろその経験者の方が成績がよい)。また、現職14年未満についても3タイプとも有意差はみられなかった。

なお、今回の分析では前述のようにインシデントのエラーのみを分析した(低リスク輸送障害経験者のエラーは分析していない)。一方、事故ランク別の分析で、インシデント経験者と低リスク輸送障害経験者のPF値に差が認められなかったことから、両者の区分は特に必要ないと考えられる。したがって今後は、インシデントを含む事故・輸送障害の全エラーを対象に同様の分析を行うことにより、PF値とエラーとの関係がより明確になると考えられる。

考 察

内田クレペリン検査の識別性については、まず事故ランク別の分析から、現職14年の運転士についてPF値と輸送障害との関係が認められなかった。一方、現職14年未満の運転士については、現職3年未満の運転士についてのみ、PF値と輸送障害との関係が示唆された。したがって、内田クレペリン検査は、運転経験の浅い時期においてヒューマンエラーに対する識別性を持つ可能性があり、逆に経験を積むと識別性が薄れる可能性がみられた。

一方、ヒューマンエラーの3つのタイプについては、スリップとミスタイクについて識別性の可能性が示唆されたが、データの統制が十分に行えないため、明確な結論は得られなかった。

また、インシデントと低リスク輸送障害ではそれらを引き起こすエラーの質に差があると想定し、両者の経験者を区分してPF値を算出したが、両者に差はみられなかった。

以下では、得られたデータや識別性についての解釈、あるいは今後の識別性評価の課題などについて、先行研究の内容を踏まえ考察する。

内田クレペリン検査の識別性に関する考察

今回の分析では現職3年未満について識別性が示唆されたが、前述のように完全無事故者と輸送障害経験者のPF値の平均値の差は、標準偏差を考慮す

るとそれほど大きくはなかった。

最近の研究で、今回の分析と同様に、事故等の経験者と無事故者のPF値の平均値の差を分析しているのは、井上・鈴木・喜岡・赤塚・重森・樋田(2008)のみである(ただし、サンプルは運転士のみではなく、車掌、指令員等、運転関係従事員全般)。その結果、有意差は確認されたが、数値は事故等の経験者群が11.6、無事故者群が10.7であり、また井上他(2008)に基づき効果量(Cohenの*d*)を算出すると0.17であった。今回の分析結果(輸送障害経験者が9.83、完全無事故者が8.92で、Cohenの*d*が0.22)と比べると、サンプルが異なるので平均値は異なるが、差は同程度である。本検査における“事故”群と“無事故”群のPF値の差のレベルを示すものといえる。

一方、これらの結果の評価には以下の点も考慮しておく必要がある。まず、本分析は、検査の合格者のみを対象に行われている。したがって、低成績者があらかじめ除かれた集団での分析結果であるという点である。また、予測対象の事故やエラーについても、その発生には環境要因など人以外の要因が影響している点にも考慮が必要である。

今後、内田クレペリン検査の識別性をより正確に評価していくためには、現在の不合格者も含めたパフォーマンス評価がもてられる。不合格者が運転業務に従事することは規程上不可能なので、スクリーニングされていない集団を対象に、運転作業に類似させた課題を実施させ、検査成績とパフォーマンス(エラーの有無等)を比較するなどの地道な研究が求められる。

識別性がみられた原因の考察

次に、識別性の原因について考察する。この点についての研究は十分に行われていないが、検査結果とヒューマンエラーとの関連(井上他, 2006)、あるいは間接的ではあるが、エラー発生につながる注意の特性との関連(喜岡他, 2003; 喜岡他, 2004)について一定の知見がある。また、検査結果と性格特性の関連についての研究もある(柏木, 1995; 瀧本, 2000; 鈴木他, 2004)。その内容にふれながら識別性の原因について考察する。

ヒューマンエラーとの関連 まず、検査成績による発生しやすいエラータイプの違いが井上他(2006)において分析されている。この研究では、前述の

Reasonとは異なる6種類からなるエラー分類を行い(エラーパターンとよばれる),各パターンのエラーを誘発しやすい実験課題の成績と内田クレペリン検査の成績との相関を求めている。内田クレペリン検査のPF値と有意な相関が認められたのは,6パターンのうち,習慣エラー(注意のそれ)のみであった。このエラーは,重森・井上・澤(2006)によると,普段とは違う作業をしなければならないのに,別の作業と並行して当該作業を実施したため,注意容量が不足し,普段どおりの行動をとってしまうエラーである。具体的な事例としては,運転操作や事故復旧作業など,求められている正しい手続きを行わなければならないときに,同時に遅延の回復や他の作業などに注意を奪われ(逸れ),目の前の作業に十分に注意が向かない場面などが相当するとされる。したがって井上他(2006)の結果から,内田クレペリン検査のPF値は,習慣的行為を抑制できる「習慣抑制力」,および注意が逸れにくいあるいは逸れても十分な注意を残すことができる「注意容量」の2つの能力と何らかの関係があると推定できる。一方,本研究の分析2ではReasonの分類を用いてエラータイプ別のPF値を分析した結果,スリップおよびミステイクについて識別性の可能性が示唆された。スリップとは,「習熟した作業において自動的なルーチンが動作しているときに,通常では発生しない変化が生じたが,注意力の資源が当該作業とは無関係な対象に投入されているために,その変化に気づかず通常通りの行動をしてしまうこと」が典型とされる(Reason & Hobbs, 2003 高野監訳 2005)。これは,上記の習慣エラー(注意のそれ)と共通要素の多いエラーといえる。また,典型的なミステイクは,「ルールに従った作業を行う際,条件の多くは当該ルールの適用条件に合致していたが,当該ルールを適用してはいけない条件があり,それを見落としたために,間違った行動をしてしまったもの」で,習慣の影響や状況の変化に気づかなかったことが影響するとされる(Reason & Hobbs, 2003 高野監訳 2005)。上記ミステイクの中には,ルールの適用条件の吟味を行う際に,他の作業に注意力の資源を奪われたために,この吟味が不十分になり,その結果として,普段はめったに生じない「当該ルールを適用してはいけない条件」の見落としが生じ,そのため普段どおりの行動をしてしまうケー

スがある。このタイプのミステイクは,前記の習慣エラー(注意のそれ)の「普段と違う作業をしなければならないのに注意容量の不足から普段どおりの行動をしてしまった」ことと内容的に類似する要素がある。したがって,詳細は今後の研究によらざるを得ないが,今回の分析が示唆する結果は井上他(2006)の知見とも部分的に符合すると考えられる。

また,今回の分析からは現職3年未満の運転士についてPF値に差がみられた。したがって,現職3年未満で,まだ十分経験が積まれていない運転士は,並行作業がある場合に注意配分が適切に行えず,習慣的な行動をとってしまう習慣エラー(注意のそれ)が発生しやすい可能性があるといえる。

注意との関連 一方,注意の特性と定型・非定型との関係については,喜岡他(2003)が,注意の配分力・集中力・持続力を測定する検査の成績と内田クレペリン検査の成績を比較している。その結果,曲線類型が非定型の人は注意の配分力や持続力が低いことを見出している。

注意の配分力については,ランダムに配列された数字を順に指していく検査(注意配分検査)により測られているため,前項のヒューマンエラーとの関係で取り上げられた注意配分とは,別個の独立な並行課題がない点で異なる。しかし一方で注意配分検査では当面の数字を探すとともに,それに続くいくつかの数字を可能な限り記憶することが有利になるとされることから(三井・福本・武井, 1968),前項のエラーとの関係で述べたPF値が低いと並行作業がある場合に注意配分が適切に行えないという部分では一定の関連があると考えられる。

注意の持続力については,喜岡他(2003)はその分析結果に基づき,内田クレペリン検査はタイムプレッシャーのかかる短期集中力よりも注意の持続力あるいは単調耐性に近い特性を測定していることに示唆があったとしている。また,喜岡他(2004)では,発動性,可変性,亢進性という内田クレペリン検査における性格3特性(日本・精神技術研究所, 1990)と上記3検査(注意の配分力,集中力,持続力)との関係を分析しており,その結果,亢進性が低いと注意の持続力が低いことを見出している。曲線類型と性格3特性との関係はこれまで明確には論じられていないが,喜岡他(2004)から,注意の持続力の低さは,亢進性の低さ(心的活動水準が継時

的に低下しやすいなど）（日本・精神技術研究所，1990）との関係も示唆される。運転士の業務は、決められたダイヤに沿って信号確認・速度調整・外乱監視などの作業を繰り返す業務であり、求められる能力としてはビジランスと集中、特に繰り返し作業や単調作業時の遂行能力が重要とされる（RSSB, 2008）ことを考えると、内田クレペリン検査の定型・非定型と注意の持続力との関係は、運転士の作業特徴と符合する。

一方、本研究では、現職3年未満の運転士でPF値と輸送障害との関係が示唆され、それ以外の経験年数では関係が認められなかった。この研究結果と上記の注意との関係については推定の域を出ないが、まず、経験が浅い運転士の場合には、自身の弱点である注意の配分力・持続力の低さが運転の質に直接影響してしまい、エラーにつながっていることが考えられる。一方、運転士は、自身の失敗経験など、運転経験を通して作業中に発生しうるエラーの防止スキルを獲得している（楠神・武田，2005）。したがって、運転士には経験を通して自身の弱点（この場合は注意の配分力・持続力）が意識され、それをカバーするためのノンテクニカルスキル（Flin, O'Connor, & Crichton, 2008）が蓄積される。それらの結果として、経験年数が長くなると検査の識別性が低下する。以上が一つの仮説として考えられる。

なお、定期検査は、先述のように何らかの理由により著しい機能低下を来した少数例を発見するためのチェック機能を有するため、本研究の結果があっても、定期検査の意味は継続すると考えられる。

その他、性格についてもレビューし、PF値がYG性格検査のAGと有意な相関がある（瀧本，2000）、PF値はビッグファイブの5因子すべてと有意な相関があるが係数が高くない（柏木，1995）等の知見は得られたが、エラーとの関連からみた本検査の測定特性を推定することはできなかった。

インシデントと輸送障害のPF値に関する考察

また、今回はじめてインシデント経験者と低リスク輸送障害経験者を区分して分析したが、全体および5つの群すべてで両者のPF値に差が認められなかった。事故・インシデントの防止が主な目的である信号保安装置等のハード対策の導入により、インシデントを惹起するエラーは異常時にシフトしてい

ると考えられるが、調査時点ではエラーの質に大きな差はないと考えられる。割合の変化は今後も続くと考えられるため、同様の分析を継続する必要がある。また、今回は、インシデントと輸送障害の比較のみで、平常時と異常時のエラーを区別することはしなかった。異常時エラーの重要性は今後も高まると考えられ、また前述のように異常時エラーの発生には情動の影響が想定されるため、本検査と情動との関係の評価も含め、異常時エラーという括りで内田クレペリン検査の識別性を分析することが別途必要である。

今後の課題

内田クレペリン検査については、測定内容や解釈の曖昧さに関して批判が多い。たとえば、村上（2005）は、プロフィールの激しい動揺を気分や情緒が不安定と解釈する点には疑問があるとし、また定型パターンや非定型パターンの指標は不安定で妥当性の高い解釈はできないのに、無理に性格を解釈する点が問題としている。また、吉田（2006）も、内田クレペリン検査が何を測定するのか心理学的にはわからないし、数々の問題があるとしているが、一方で、実用の立場からは、事故を起こす確率にわずかの違いがあれば安全策を取るのが経営の責任なので、検査に重きを置く根拠はあるとも述べている。

したがって今後は、検査の実施の根拠をより明確にしていくため、検査成績が事故と関連をもつ程度や根拠、あるいは測定特性等の明確化に向けて引き続き心理学的な分析を行っていく必要がある。その際には、運転士の「経験年数」という切り口での分析が必要と考えられる。今回の分析では、現職3年未満のみ検査成績と輸送障害との関連が認められるなど、経験年数により検査の識別性が異なる可能性が示されたからである。

また、本研究では、輸送障害等と検査結果の関連をみた上で、識別性の根拠についてエラーや注意との関係から考察したが、今後は、より統制のとれたデータの収集に加え、エラーのメカニズムあるいは注意以外の認知特性との関連分析が重要である。また、性格との関係やPF値とは基本的に関係ないとされる知能（瀧本，2000）などもその確認の対象となる。

一方、内田クレペリン検査の識別性評価に当たって、今回は、運転適性検査の前提からPF値に経年

変化がない前提で分析を行ったが、今後は、より厳格な評価に向け、縦断的データによるPF値の経年変化の有無についても評価していく必要がある。

また、識別性の評価に当たっては、PF値や曲線類型を用いている限り、分析に限界がある。本検査には、激しい動揺、後期作業量の下落、後期初頭の出不足など、さまざまな非定型特徴がある。一方でPF値あるいは曲線類型は、それらの非定型特徴の程度を総合的に表す指標であるため、その数値からはどの非定型特徴の影響が大きいかを読み取れない。したがって、各非定型特徴について、それぞれがエラーや認知特性にどう関係するかを分析することにより、検査の測定特性に関する知見が深まると考えられる。実際、自動車分野では長塚(1985)が、事故群と無事故群の検査成績の違いをもとに、事故の発生と初頭努力、休憩効果、動揺率との関係を明らかにしている。また、各分作業量のデータからこれらの非定型特徴の程度を個別に定量化する方法も提案されている(小笠原, 1985)。したがって、鉄道において個々の非定型要素に着目することが今後の有力なアプローチの1つといえる。それにより、各非定型特徴と特定のエラーの発生メカニズムあるいは認知特性との関連が推定でき、上記の課題検討の足掛かりになると考えられる。

引用文献

- Flin, R., O'Connor, P., & Crichton, M. 2008 *Safety at the Sharp End: A Guide to Non-Technical Skills*. London: Ashgate.
- Gibson, W. H., Smith, S., Lowe, E., Mills, A. M., Morse, G., & Carpenter, S. 2013 Incident Factor Classification System. In Dadashi, N., Scott, A., Wilson, J. R., & Mills, A. M. (Eds.), *Rail human factors*. London: Taylor & Francis, pp. 653-658.
- 細田 聡・井上枝一郎 2000 緊急事態での人間行動の特徴に関する一考察 労働科学, 76, 519-538.
- 池田敏久・大嶽ヒサ 1974 動力車乗務員の誤扱い事故防止のためのバックアップシステムの考え方 鉄道労働科学研究資料, 74-11.
- 井上貴文・重森雅嘉・喜岡恵子・赤塚 肇・宮地由芽子 2006 新しい運転適性検査項目の提案 鉄道総研報告, 20, 5-10.
- 井上貴文・鈴木浩明・喜岡恵子・赤塚 肇・重森雅嘉・樋田 航 2008 新しい運転適性検査体系 鉄道総研報告, 22, 5-10.
- 柏木繁男 1964 内田クレペリン検査の信頼性と妥当性の客観的手法による検討 心理学研究, 35, 93-95.
- 柏木繁男 1975 内田クレペリン検査における解析的評価法 金子書房.
- 柏木繁男 1995 性格特性5因子モデル(FFM)による内田クレペリンテストの評価について 心理学研究, 66, 24-32.
- 喜岡恵子 2000 運転適性検査制度の現状と今後の心理適性検査研究の課題 鉄道総研報告, 14, 1-6.
- 喜岡恵子・鈴木綾子・宮地由芽子 2004 鉄道における運転適性検査の研究(11)—内田クレペリン検査における3特性と注意力— 日本教育心理学会総会発表論文集, 46, 472.
- 喜岡恵子・宮地由芽子・山内香奈 2003 鉄道における運転適性検査の研究(4)—作業性検査と注意力(注意配分・集中・持続)との関連性— 日本教育心理学会総会発表論文集, 45, 187.
- 楠神 健・武田祐一 2005 ベテラン運転士が持つエラー防止スキルの分析—リスクモニタリング訓練手法の研究(1)— 産業・組織心理学会第21回大会発表論文集, 79-82.
- 楠神 健 2013 異常時パフォーマンスが予測可能な多重選択反応検査の作成 日本テスト学会誌, 9, 57-68.
- 三隅二不二・白樫三四郎・安藤延男・黒川正流 1961 内田クレペリン精神検査の妥当性に関する研究 教育・社会心理学研究, 2, 139-158.
- 三井大相・福本昌子・武井敏一 1968 注意の研究(その1) 鉄道労働科学, 22, 1-16.
- 長塚康弘 1985 事故多発運転者の作業特性についての研究—内田クレペリン精神作業検査の妥当性の検討を通じて— 交通心理学研究, 1, 25-35.
- 日本・精神技術研究所(編) 1975 内田クレペリン精神検査・基礎テキスト 日本・精神技術研究所.
- 日本・精神技術研究所(編) 1990 内田クレペリン精神検査データブック 金子書房.
- 村上宣寛 2005 「心理テスト」はウソでした. 日経BP社.
- 大庭幸穂 1981 国鉄における適性検査の歴史(その1) 鉄道労働科学研究資料, 80-23.
- 小笠原春彦 1985 内田クレペリン検査の数量判定法 鉄道労働科学研究資料, 85-20.
- Rasmussen, J. 1983 Skills, rules, and knowledge; Signals, signs and symbols, and other distinctions in human performance model. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13, 257-266.
- Reason, J. 1990 *Human Error*. Cambridge: Cambridge University Press.
- Reason, J., & Hobbs, A. (2003). *Managing Maintenance Error: A Practical Guide*. London: Ashgate.
- (リーズン, J.・ポップズ, A. 高野研一(監訳) (2005). 保守事故 日科技連出版社)

RSSB 2008 *Rail Industry Standard for Train Driver Selection*.
London: RSSB.

重森雅嘉・井上貴文・澤 貢 2006 ヒューマンエラー発生傾向を測定するための模擬課題の開発 鉄道総研報告, 20, 11-16.

相馬紀公 1949 内田クレペリン精神反応検査の数量的取扱 日本・精神技術研究所.

鈴木綾子・喜岡恵子・宮地由芽子 2004 鉄道における運転適性検査の研究 (12)—YG 性格検査による内田クレペリン検査の測定内容検討の試み— 日本教育心理学会総会発表論文集, 46, 473.

瀧本孝雄 2000 内田クレペリン精神検査の妥当性に関する心理学的研究 (6) *Mathesis Universalis* (獨協大学諸学研究), 1, 450-496.

瀧本孝雄 2013 内田クレペリン精神検査の妥当性に関する総合的研究 *Mathesis Universalis* (獨協大学諸学研究) 14, 9-35.

鉄道労働科学研究所名古屋駐在 1970 運転事故者の特性についての一考察 鉄道労働科学研究資料, 70-20.

鉄道労働科学研究所労働心理研究室・人間工学研究室・心理適性管理室 1982 誤出発事故防止の労働科学的研究 鉄道労働科学研究資料, 82-2.

鉄道労働科学研究所心理適性管理室 1980 心理適性検査研究の変遷と将来像 鉄道労働科学研究資料, 79-52.

鉄道労働科学研究所心理適性管理室 1987 心理適性検査研究の変遷と将来像 (続) 鉄道労働科学研究資料, 86-37.

鶴田正一 1950 運転保安科学〈運転事故防止の心理学的問題〉 鉄研資料, 19.

数原 晃 1986 事故者の性格・態度特性について (1) 鉄道労働科学, 40, 165-178.

数原 晃 1990 責任事故者の運転適性検査データの分析 鉄道総研報告, 4, 9-17.

米山信三・宇賀神 博 1986 場内信号冒進事故の分析と対策 鉄道労働科学研究資料, 85-37.

吉田信彌 2006 事故と心理 中央公論新社.

(受稿：2014.6.30；受理：2015.5.25)