

非定常時・異常時 パフォーマンスを測定可能な 多重選択反応検査の開発



青沼 新一*



静山 弘美*



高田 雄一郎*



楠神 健*

現在の鉄道の運転適性検査体系は、1949年の制定以来大きな見直しを受けることなく今日に至っている。その間、運転環境は大きく変化し、要注意なヒューマンエラーの内容も非定常時・異常時にシフトしてきた。このため、既存の運転適性検査の事故やエラーに対する識別力を考慮しながら、今後、重要になると想定される非定常時・異常時パフォーマンスの予測が可能な適性検査を複数の選択反応（色・形・音）を組み合わせた多重選択反応検査として開発した。また、その妥当性（予測可能性）を、非定常時・異常時パフォーマンスに加え各種エラーおよび実際の運転作業における取扱い誤り事象に対して確認した。

●キーワード：運転適性検査、多重選択反応検査、心理適性検査、異常時パフォーマンス、ヒューマンエラー

1. 研究の背景と目的

運転適性検査とは、国土交通省の「鉄道に関する技術上の基準を定める省令」の中に「列車等の運転に直接関係する作業を行う係員が作業を行うのに必要な適性、知識および技能を保有していることを確かめること」と定められており、身体機能検査ならびに精神機能検査を定期的に受検することが定められている。また「動力車操縦者運転免許に関する省令」においても動力車操縦者運転免許を取得する際に、クレバリン検査および反応速度検査を受検することが定められている。これを踏まえJR東日本では運転士への登用試験の際に臨時検査を、運転取扱作業に従事する社員に対しては3年に一度の定期検査をそれぞれ実施している。

運転適性検査体系は1949年の制定以来大きな見直しを受けることなく今日に至っている。検査体系を制定した1949年当時は、運転士が乗務する車両は蒸気機関車が中心であり、列車を運転すること自体が熟練を要する作業であった。一方、現在乗務する車両はそのほとんどが電車・気動車になったため、列車の加減速のためのハンドル操作も比較的容易になり、車両状態等も常にモニタで監視可能になってきた。また、ATS・ATCなどの保安装置の整備に伴い、運転士の「うっかりミス」が即重大事故につながるリスクも大幅に減少している。このように運転環境は大幅に変化しているにもかかわらず、現行の運

転適性検査の検査成績と鉄道運転事故との関係は十分に明らかになっているわけではない。

そこで、当社では2001年度下期～2003年度にかけて『運転適性検査体系の今後のあり方に関する検討会』（以下、「第一期検討会」と呼ぶ）を設け、運転士を対象に現行運転適性検査の識別力を明確にすると共に、現在および今後の運転環境に適合した新しい運転適性検査のあり方について検討を行った。

その結果、一部識別力の低い検査も見られたが、検査体系全体としては一定の妥当性が確認された。しかし、前述のように平常時のうっかりミスよりも非定常時・異常時の判断エラーの方が、鉄道運転事故につながるリスクが高いことも確認された。

そのため、第一期検討会では、以下のことが提言された。「鉄道運転事故を未然に防止していくためには、検討会において非定常時・異常時パフォーマンスに一定の識別力が確認された現行の機敏性検査（省令でいう反応速度検査）を発展させ、識別力を強化するとともに、識別性検査もその新検査で代用していく。また、新検査は、機敏性検査と同様、「選択反応をベースとした検査」とする。」

そこで、本研究では、2004～2005年度に「運転適性検査の開発・利用に関する検討会」を設け、全職種に適用可能な非定常時・異常時パフォーマンスの識別力（予測力）を強化した新検査を試作し、以下の観点から運転適性検

査としての妥当性の確認を行った。

- (1) 非定常時・異常時パフォーマンスに対する予測力評価
- (2) 各種のヒューマンエラーのパターンに対する予測力評価（財団法人鉄道総合技術研究所に依頼）
- (3) 取扱い誤り事象発生に対する予測力評価

2. 非定常時・異常時エラーと関係する心理特性

2.1 非定常時・異常時の行動モデルの作成

非定常時・異常時のパフォーマンスに係わる心理特性を検討するため、文献 1)、2) をベースに、非定常時・異常時の行動モデルを検討し、以下のような簡単なモデルを作成した（図1）。

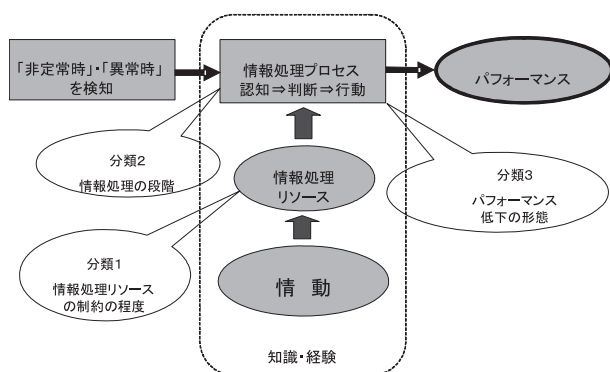


図1 非定常時・異常時の行動モデル

作業者は、検知をきっかけに、既存の知識・経験を駆使しながら、認知・判断・行動の一連の情報処理を行い、一定のパフォーマンスを発揮すると考えられる。この際、作業者が遭遇した状況を「非定常時・異常時」と認知すると、恐怖・不安・焦りなどの情動が生じ、それが情報処理のリソースに制約を加え、パフォーマンスに影響を与える、と考えるモデルである。

2.2 非定常時・異常時エラー等の分析

このモデルをベースに、鉄道における非定常時・異常時エラーを分析し、非定常時・異常時のパフォーマンスに関連する心理特性を検討した。具体的には、上記モデルの中の「情報処理リソースの制約の程度」「エラーが発生しやすい情報処理の段階」および「パフォーマンス低下の形態」について、1988年から2001年までに発生したリスクの高い非定常時・異常時エラー75件について分析し

た。エラーの内訳を表1に示す。

表1 非定常時・異常時エラー内訳

事象	件数
防護無線	23
独断退行・運転継続	18
信号・合図関係	17
保安装置関係	7
異常時の処置	7
その他（停止位置・ドア・点検）	3
計	75

非定常時・異常時の情報処理の特徴を上記の3つの軸から分析した結果、検査課題の設定に際しては、以下の点への留意が重要と考えられた。

(1) 情報処理リソースの制約の程度

真っ白、空白、放心状態になった事例はなかった。しかし、特に危険度の高い事象（たとえば、列車脱線時に列車防護の措置が適切に取れない）については、その可能性をゼロと想定することはできないと考えられた。そのため、時間内に、速く正確に作業させることを求めるなど、課題遂行に対するプレッシャーを与える課題を作成し、その中でパフォーマンスをみる必要がある。

(2) エラーが発生しやすい情報処理の段階

情報処理の段階では、「判断」でのエラーが最も多く、「認知」が次に多かった。そのため新検査においては、「状況・原因に応じた行動を判断する際のエラー」や「状況・原因を認知する段階でのエラー」を想定した課題設定が必要である。

(3) パフォーマンス低下の形態

発生したエラーを見ると、非定常時・異常時に情報処理リソースを奪われる中、「普段の行動パターンに影響されて行動が特定の方向に引きずられるタイプ」よりも、「行動できなかった、あるいは行動できたが、確認忘れ、手順漏れ、判断誤りが発生した、あるいは、状況・原因自体が十分認識できなかったタイプ」の方が多かった。特にリスクの高いエラーは後者が主体であったため、検査では主にこのエラーを模擬できることが望ましい。

3. 多重選択反応検査の選定理由

3.1 新検査の開発コンセプト

2章の分析をまとめると、鉄道における非定常時・異常時パフォーマンスを規定するのは、「情動の影響により情報処理リソースが制約される中で適切に認知・判断・行動できる能力」といえる。また、この能力は、たとえば「定常時の情報処理リソース」「情動の発生のしやすさ」「情動が情報処理リソースに及ぼす影響」など多くの要素に分解できるが、検査として最も予測したいのは、これらの影響が総合された最終的なパフォーマンスであるとされた。

3.2 多重選択反応検査の定義

上記のコンセプトを実現するために、多くの種類の検査の中から多重選択反応検査を選定したが、その理由を述べる前にまず、「多重選択反応検査」の定義を行う。多重選択反応検査とは、「選択反応の系列を複数組み合わせた課題を設定し、任意の系列の任意の刺激をランダムに提示し、受検者にその刺激に対応した反応をできるだけ速く正確に繰り返すことをもとめる検査」とする。系列の例としては、色の選択、形の選択、音の選択などが考えられる。

3.3 多重選択反応検査の選定理由

非定常時・異常時パフォーマンスの予測が可能な検査として、多重選択反応検査を選定した理由を以下にまとめる。

- (1) 1系列の選択反応である機敏性検査においてすでに非定常時・異常時エラーに対する一定の識別力が確認されているが、多重選択反応検査では、選択反応を複数組み合わせることによって、より複雑かつ難易度の高い課題設定が可能である。したがって、不安やあせり等の情動により、リソースが制約された場面での作業をより模擬できると推定できる。
- (2) 多重選択反応検査では、反応が一定以上遅い場合に、視覚あるいは聴覚刺激を用いてフィードバックを与えることにより、受検者にタイム・プレッシャーを与え、より非定常時・異常時に近い場면을模擬できていると考えられる。

- (3) 機敏性検査での1系列の選択反応は、どちらかといえば単純な情報処理であり、「一つの作業の流れに沿って淡々とルールに従った行動を行う定常時の作業」に近いといえる。一方、多重選択反応検査での複数系列の選択反応は、より複雑な情報処理が必要であり、「短い時間の中で、いろいろなことを同時並行的に処理しなければならない非定常時・異常時の作業系列」により近いと考えられる。
- (4) 多重選択反応検査の課題は、「正確な認知」がベースになっている点はもちろんであるが、複数の系列の表示情報と操作情報を理解した上で、それらの情報を適切に関連づけ、組み合わせることで正しい反応ができる点で、「認知」のパフォーマンスに加え、「判断」のパフォーマンスも評価できると考えられる。また、一定のルールのもとで、不慣れな複数の課題を、スピーディかつ正確に処理することが求められる非定常時・異常時の作業に性格的に類似した課題設定が可能と考えられる。

以上の理由から、新検査は多重選択反応検査をベースに開発することにした。

4. 多重選択反応検査の仕様の検討

これまでの調査・分析を踏まえて、多重選択反応検査の仕様を検討した。具体的には、選択反応課題の内容、刺激の提示方法、反応が遅い場合のフィードバックの与え方、検査の実施方法、教示、練習の内容、記録データ（検査指標）などについて検討した。

その結果、検査は、色・形・音の3つの選択反応課題から構成することにした。また、検査試作後、研究開発センター内の社員と実際の運転士に検査を実施し、課題の難易度のチューニングや、練習内容・教示等の改善を繰り返した。

以下に、改善を経て確定した多重選択反応検査の内容、実施方法等を説明する。

(1) 多重選択反応検査の内容

- ・検査は、パソコン、キーボード、ヘッドホンで構成される。
- ・選択課題としては、「色」「形」「音」の3種類の系列を選定した。複数のモダリティに対応した刺激

を用意したのは、非定常時・異常時の作業環境の模擬を意識したためである。各系列の刺激にはそれぞれ3種類の刺激が用意されている。検査は、任意の系列の任意の刺激（3×3の9種類）がランダムにディスプレイあるいはヘッドホンから提示され、受検者はその刺激に対応した反応（キーボードの対応するキーを押す）をできるだけ速く正確に繰り返すものである。視覚刺激は、画面上の6箇所の任意の位置に提示する。

- ・受検者に対して、反応が遅い場合には、「キーを押してください」という表示を画面下に提示する。また、提示されているのが視覚刺激の場合には、音アラームを合わせて出す。「反応が遅い」と判断する基準は、受検者各人の反応速度から設定される。
- ・刺激に対する反応（キー押し）が正解の場合には、即座に次の刺激が提示される。不正解の場合には、刺激が切り替わらない。この点は機敏性検査と同じである。
- ・検査所要時間 全体：4分強、本検査：60秒
検査の詳細を図2に示す。

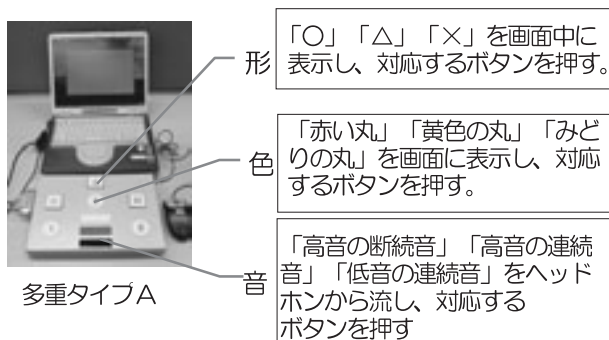


図2 多重選択反応検査の内容

(2) 検査の実施方法

検査の実施方法を以下に簡単にまとめる。

- ・受検者はヘッドホンをつける。
- ・目視点は画面中央とし、左手は①と書いてあるキー、右手は③と書いてあるキーの手前に置く。
- ・画面に色・形の刺激が、ヘッドホンからは音の刺激がランダムに提示される。
- ・受検者は、できるだけ速かつ正確に、提示された刺激に応じたキーを押す。
- ・反応が遅い場合には、前述のようにその旨を画面等

に表示する。

- ・練習後、本検査を60秒実施する。

なお、以下では上記の仕様で作成された検査のことを、多重選択反応検査ということにする。

5. 非定常時・異常時パフォーマンスの予測力に関する評価

5.1 目的

多重選択反応検査が持つ非定常時・異常時パフォーマンスに対する予測力を現行の機敏性検査との比較を通して明らかにする。

5.2 評価方法

総合訓練センターにおいて、運転士190名に多重選択反応検査および機敏性検査を受検してもらった。一方、当該運転士の素質面からみた異常時の強さを「訓練センター講師」（異常時のシミュレータ訓練3）を担当「職場の上司」「運転士本人」の三者に評価してもらった。

5.3 評価結果

「多重選択反応検査の成績」と「三者それぞれの評価・三者の平均値」との相関係数を算出すると、「三者の平均値」「訓練センター講師」「職場の上司」「運転士本人」の順になった。

ここでは、非定常時・異常時パフォーマンスの指標として、相関の最も高かった三者平均を用いて、多重選択反応検査と機敏性検査の予測力を比較する。比較した検査指標は、いずれも正答数と誤答数（現在の機敏性検査で合否判定に用いられている指標）である。また、素質面からみた異常時の強さは、以下の6つの観点から5段階で評価してもらったが、非定常時・異常時パフォーマンスの評価には最後の総合評価を用いた。

- ①異常時でも、状況をすぐに正しく理解できる
- ②緊急動作（列車防護など）をためらわずに行える
- ③指令への連絡・確認など、大事な手順は漏れなく実施できる
- ④危険かもしれないと思えば、しっかり列車を止めておける
- ⑤パニックにならず、落ち着いて判断・行動できる
- ⑥総合的にみて異常時に強い

以下の図3は、両検査の正答数について、検査の予測力の評価に用いられることの多い方法である、検査成績を

上位・中位・下位に3分割したうえで「非定常時・異常時の強さ」の評定値を比較したものである（3分割は、人数ベースで各層に1/3が入るように分割）。

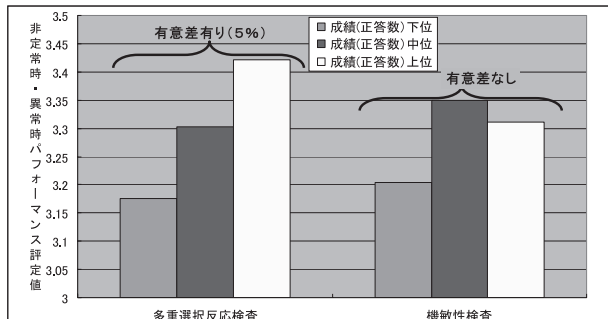


図3 非定常時・異常時パフォーマンスの予測力評価

多重選択反応検査は、成績が上がるほど非定常時・異常時パフォーマンスの評定値が高くなっており、分散分析の結果、有意差が認められた（ $p<0.05$ ：危険率5%で有意差が認められたことを示す）。一方、機敏性検査については、そのような傾向は見られなかった。なお、誤答数については、多重選択反応検査、機敏性検査とも成績が上がる（誤答数が少ない）ほど非定常時・異常時パフォーマンスの評定値が高くなる傾向がみられたが、いずれも有意差は認められなかった。

非定常時・異常時の強さを表すその他の項目については、両検査とも「③指令への連絡・確認など、大事な手順は漏れなく実施できる」「④危険かもしれないと思えば、しっかり列車を止めておける」の相関が相対的に高かった。これらに次ぐのが「②緊急動作（列車防護など）をためらわずに行える」であった。

6. 各種エラー発生に対する予測力の評価

6.1 目的

多重選択反応検査のヒューマンエラーに対する予測力の評価が（財）鉄道総合技術研究所（以下鉄道総研）により行われている。ここでは、文献4）に基づき、評価結果を簡単にまとめる。

6.2 評価方法

鉄道総研のエラーモデルによる6つのエラーパターン（12指標）に対する予測力の評価が行われた。具体的には、一般の被験者79名（19～24歳の若年男性30名、19～22歳の

若年女性29名、55～69歳の高年齢男性20名）に対して、6つのエラー模擬実験を実施すると共に、多重選択反応検査および機敏性検査を受検してもらい、その相関から予測力を評価した。

6.3 評価結果

6つのエラーパターンに関する指標と検査成績との相関を表2に示す。

表2 エラーパターンと検査指標との相関係数

	多重正答数	機敏性正答数
効率エラー(ぼんやり)	-0.52**	-0.52*
効率エラー(注意逸れ)	-0.35**	-0.23
習慣エラー(ぼんやり)	-0.35**	-0.27*
習慣エラー(注意逸れ)	-0.34**	-0.34*
失念エラー(ぼんやり)	-0.31**	-0.19
失念エラー(注意逸れ)	-0.19	-0.18

+: $p<0.10$ *: $p<0.05$ **: $p<0.01$

※文献4) p9 表5「エラー模擬試験と各検査との相関係数」より一部抜粋

負の相関係数が高いほど、検査のエラー予測力が高いことを示す。表より、多重選択反応検査は5タイプのエラーに予測力があるといえ、機敏性検査の3種類より多い。また、相関係数も全般的に多重選択反応検査の方が高いことがわかる。

また、個々のエラータイプの説明は、文献5）に譲るが、非定常時・異常時パフォーマンスの予測を目指してデザインした多重選択反応検査であったが、ぼんやり型のエラーについても、広く予測力が認められた。

7. 取扱い誤り事象発生に対する予測力の評価

7.1 目的

多重選択反応検査の取扱い誤り事象（受検者の実作業場面での取扱い誤り）の発生に対する予測力との関係を機敏性検査との比較を通して評価した。

7.2 評価方法

運転士194名に対し、多重選択反応検査および機敏性検査を受検してもらい、その受検者の過去6年間に発生した取扱い誤り事象の有無と両検査の結果との対応をみた。なお、194名中、取扱い誤り事象の経験者は62名であった。

7.3 評価結果

取扱い誤り事象を経験した受検者と未経験の受検者のそれぞれについて、多重選択反応検査と機敏性検査の正答数、誤答数の平均値を算出した。両検査の正答数についての評価結果を図4に示す。

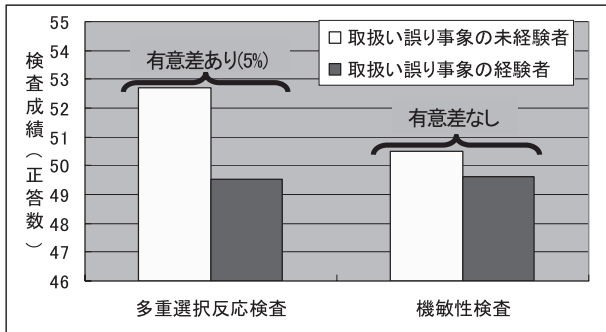


図4 取扱い誤り事象に関する評価結果

図より、多重選択反応検査、機敏性検査の正答数とも取扱い誤り事象の未経験者の方が、経験者よりも成績がよいことがわかるが、両者に平均値の差の検定 (t検定) を適用したところ、多重選択反応検査の正答数においてのみ、未経験者と経験者の成績の間に有意差が認められた ($p<0.05$)。

一方、誤答数については、両検査とも、取扱い誤り事象の未経験者の方が経験者よりも誤答数が多いという想定とは反対の結果になった。ただし、両検査とも有意差は認められなかった。

8. 多重選択反応検査の年齢別検査成績の分析

8.1 目的

多重選択反応検査と機敏性検査の年齢別検査成績を比較し、加齢による傾向の変化を分析した。

8.2 分析方法

5章において使用した運転士190名の多重選択反応検査および機敏性検査の正答数・誤答数を、受検者の年齢別 (10歳単位) に分類し、その成績を比較した。

8.3 分析結果

多重選択反応検査と機敏性検査の正答数について、年代ごとの検査成績の平均を図5に、標準偏差を図6に示す。

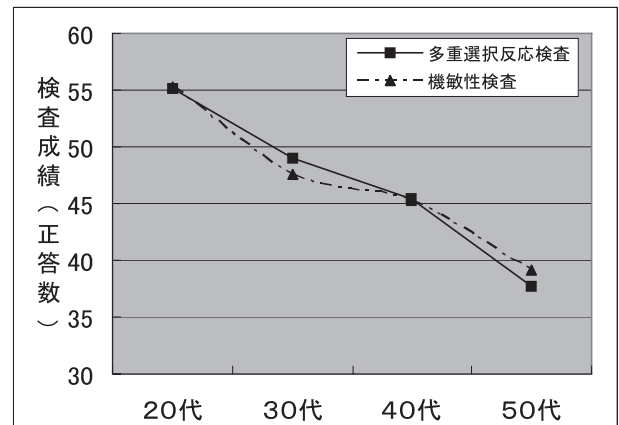


図5 多重選択反応検査（正答数）と機敏性検査（正答数）の年代別検査成績（平均）

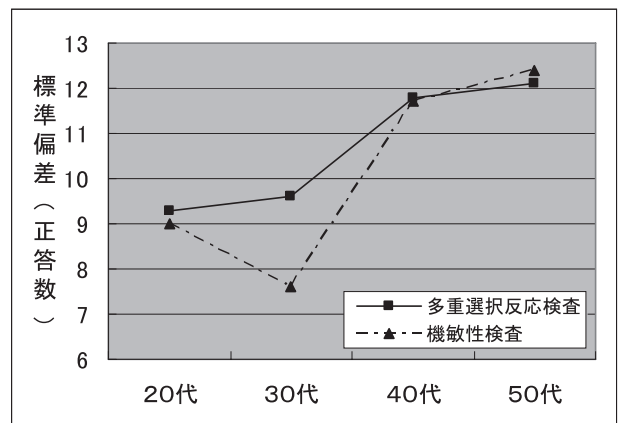


図6 多重選択反応検査（正答数）と機敏性検査（正答数）の年代別検査成績（標準偏差）

図より、検査成績は多重選択反応検査、機敏性検査とも、加齢に伴い低下がみられるが、その傾向はきわめて類似している。また、標準偏差についても30代を除けば類似の傾向がみられる。

参考までに両検査の誤答数についても同様の分析を行ったが、その結果は、平均値・標準偏差ともやはり正答数と同様、両検査で類似の傾向がみられた。

9. 多重選択反応検査と機敏性検査の指標間相関

9.1 目的

多重選択反応検査と機敏性検査の指標間の相関をみることにより、両検査の測定領域の異同について評価する。

9.2 分析方法

5章において使用した運転士190名の多重選択反応検査および機敏性検査の正答数、誤答数について、指標間の

相関係数を算出した。

9.3 分析結果

多重選択反応検査の正答数・誤答数と機敏性検査の正答数・誤答数との相関を表3に示す。

表3 多重選択反応検査と機敏性検査の指標間相関

		多重選択反応検査	
		正答数	誤答数
機敏性検査	正答数	0.606**	-0.042
	誤答数	-0.162*	0.383**

*: $p < 0.05$ **: $p < 0.01$

多重選択反応検査の正答数と機敏性検査の正答数には高い相関が認められた。一方、多重選択反応検査の誤答数は機敏性検査の誤答数と一定の相関がみられた。以上のことから、多重選択反応検査は機敏性検査と同種の検査であることがうかがわれた。

10. 多重選択反応検査の測定特性に関する考察

これまでの分析結果を踏まえ、多重選択反応検査の測定特性について考察した。

- (1) 多重選択反応検査では、機敏性検査では1系列であった選択反応の系列を3つ用意し、より複雑かつ難易度の高い課題設定を行うことにより、「系列ごとに決められたルールを理解した上で、ランダムに提示される刺激に対してルールに沿ってできるだけ速くかつ正確に反応できる力をみる検査」ということができる。
- (2) 鉄道総研の分類で言えば、①「ぼんやり」による習慣エラー・効率エラー・失念エラーおよび②「注意の逸れ」による習慣エラー・効率エラーに予測力が認められた。非定常時・異常時パフォーマンスの予測を目指してデザインした検査であったが、ぼんやり型のエラーについても、広く識別力が認められた。
- (3) 多重選択反応検査は、複雑な課題や遅延のフィードバックの提示などにより、不安やあせり等の情動により、リソースが制約される場面の作業を模擬するとともに、「短い時間の中で、いろいろなことを同時並行的に処理しなければならない非定常時・異常時の作業系列」に近い課題設定になっていると考えられる。

- (4) 多重選択反応検査は、当初のコンセプトでは、非定常時・異常時パフォーマンスの予測を目指した検査であり、測定すべき特性は「情動の影響により情報処理リソースが制約される中で適切に認知・判断・行動できる能力」とされていたが、今回の分析からは、非定常時・異常時以外のエラーについても広く予測力が確認された。したがって、多重選択反応検査は、異常時・平常時を問わず、「複数の課題をルールに沿って遂行する作業において、与えられた刺激の知覚からルールに沿った行動の出力までの一連の情報処理および行動制御を速くかつ正確に行う能力をみる検査」とまとめることができる。

11. 今後の展開

今後、多重選択反応検査は、国土交通省の「運転士の資質向上検討委員会」の場で、更なる評価を受けることになる。上記委員会では、現行検査体系の評価や新検査体系のあり方などが検討されている。今後は、さらにデータのフィールドを鉄道事業者一般に広げて、多重選択反応検査の運転適性検査としての妥当性が評価されることになる。

【謝辞】

「運転適性検査の開発・利用に関する検討会」にご出席いただき、本研究に対し貴重なご教授や多くのアドバイスをいただきました新潟大学名誉教授の長塚康弘先生、科学警察研究所交通部交通安全研究室長の西田泰先生、名古屋大学教授の野口裕之先生、追手門大学教授の藤本忠明先生、元文教大学教授の渡邊忠先生に対し、厚く御礼申し上げます。

参考文献

- 1) 細田聡、井上枝一郎：緊急事態での人間行動の特徴に関する一考察，労働科学，Vol.76, No.12, pp.519-538, 2000
- 2) 池田謙一：緊急時の情報処理，東京大学出版会，1986
- 3) 佐藤雅文，香西恵介，蔵谷正人，葛城崇：事故予防型乗務員訓練シミュレータの開発，運転協会誌，Vol.42, No.8, pp.38-41, 2000
- 4) 井上貴文，重森雅嘉，喜岡恵子，赤塚 肇，宮地 由芽子：新しい運転適性検査項目の提案，鉄道総研報告，vol.20, No.3, pp.5-10, 2006
- 5) 重森雅嘉，井上貴文，澤貢：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，vol.20, No.3, pp.11-16, 2006