

労働価値観測定尺度短縮版の項目特性に関する研究： 項目反応理論に基づく検討

An examination of item characteristics of the Work Values Scale (short version): Based on Item Response Theory

江口 圭一
Keiichi Eguchi

要 約

労働価値観測定尺度短縮版は、古典的テスト理論に基づいて、開発が進められてきた。しかしながら、古典的テスト理論によって算出される統計量には、回答者の反応に依存するという限界がある。それに対して、項目反応理論は回答者の反応から独立して統計量を推定することができる。本研究では、2パラメタ・ロジスティック・モデルにおいて、労働価値観測定尺度短縮版の下位尺度ごとに識別力パラメタと困難度パラメタが推定された。識別力パラメタは $a = 1.230 \sim 3.531$ 、困難度パラメタは $b = -2.095 \sim 0.768$ であった。したがって、いずれのパラメタも適切な値であり、一般的な労働者を対象とした有用な尺度であることが示唆された。

キーワード：労働価値観 (work values)、労働価値観測定尺度 (Work Values Scale)、古典的テスト理論 (Classical Testing Theory)、項目反応理論 (Item Response Theory)、職業生活への適応 (adjustment to working life)

目 的

研究の背景

近年、過剰な労働ストレスによる心身の疾患や過労死、過労自殺の増加に加えて、若年層における早期離職者やフリーター、ニートの増加など、職業生活に関する様々な不適應問題への対応が社会的な課題として認識されている(江口・戸梶, 2004, 2005)。これらの問題が拡大してきた背景には、社会的・経済的要因など多様な要因が存在していると考えられるが、その一因として、労働者個々の労働価値観に着目した研究が行われている。労働価値観とは、職業生活における労働者の認知や行動に影響を及ぼす要因とされ(Adkins & Naumann, 2001; Chiu & Kosinski, 1995; 大山, 1990)、「個々人が職業生活の目的として重要であると考えられる要因」と定義される(江口・戸梶, 2005)。職業生活において労働者は、個々人の労働価値観に基づいて、自らが置かれた状況を認知し、どのような行動を選択するかを決定していると考えられる。したがって、労働価値観を測定することで、職業生活における様々な不適應問題の

原因や構造を理解し、対策検討の一助とすることができるだけでなく、将来的に不適應状態に陥りやすい人たちを早めにチェックすることも可能になるであろう。さらに、適応的な労働価値観を持つ人たちに対して、普段それほど意識する機会がない自分の労働価値観について振り返ってみることで、自らの職業生活や人生について、改めて考える機会を提供することができると思われる(江口・戸梶, 2010)。

このように、労働者の職業生活における認知や行動を理解する上でも、職業生活への適応のための支援方策を検討する上でも、労働価値観を測定することには大きな意義があると考えられる。しかしながら、これまでに開発された労働価値観、あるいはそれに類似した概念を測定する尺度は、主に若年層の職業選択場面を想定した研究で用いられてきたという背景があることから、職業興味検査に近い内容のものが多いだけでなく、測定ツールとしての信頼性や妥当性が十分に検討された尺度が少ないことも指摘される(江口・戸梶, 2005)。

そのため、これらの問題点を解消した新たな労

労働価値観測定尺度の開発が行われた（江口・戸梶, 2010）。はじめに、理論的検討を踏まえて、労働価値観の下位概念として、(1) 職業生活において、自分の能力を活用することや、それを通して達成感を得ることなど、内在的な要因を労働の目的として重要視する「内的価値志向」、(2) 自分の仕事に対する高い評価を得ることや十分な収入を得ることなど、外在的な要因を労働の目的として重要視する「外的価値志向」、(3) 働くことを通して社会や組織、同僚に対して貢献することを労働の目的として重要視する「愛他的価値志向」が想定された。次に、これら3つの下位概念に相当する質問項目が先行研究から収集され、それらを用いた質問紙調査が実施された。探索的因子分析の結果、「自己の成長」、「達成感」、「社会的評価」、「経済的報酬」、「社会への貢献」、「同僚への貢献」、「所属組織への貢献」の7因子、38項目が抽出された（江口・戸梶, 2010）。この尺度について、内の一貫性（江口・戸梶, 2006a, 2006b, 2006c, 2007a, 2007b）、安定性（江口・戸梶, 2009c）、構成概念妥当性（江口・戸梶, 2006b, 2007c）、因子的妥当性（江口・戸梶, 2007a）の検証がなされている。また、この尺度を基に、短縮版（江口・戸梶, 2009a）の開発も行われている。

古典的テスト理論の限界

これらの一連の研究から、新たに開発された労働価値観測定尺度の信頼性と妥当性が概ね満足できるレベルであることが確認された。しかしながら、調査対象者がある地域を中心としたサンプルである点については、再検討する必要性が指摘される（江口・戸梶, 2010）。人の移動が激しくなった現代社会では、祖父江（1971）が指摘するような地域性や県民性といったその地域特有のもののお考え方や価値観は小さくなり、全国で均一化・画一化が進んでいると考えられる。その一方で、失業率や有効求人倍率は地域による格差が非常に大きく（厚生労働省, 2006）、そのような社会的・経済的環境の相違が地域間での労働価値観の相違につながっている可能性もある。したがって、労働価値観測定尺度についての一連の検討結果に関して、他地域でのサンプルでも同様の結果が得られるかは不明である。

このようなデータのサンプリングに関する問題は、本尺度に特有な問題ではなく、いわゆる古典的テスト理論に基づいて開発された測定尺度やテストに共通する問題である。古典的テスト理論では、母集団（例えば、日本の労働者全体）を想定し、無作為抽出したサンプルから推測統計的な分析や考察が行われる。サンプルの抽出には、多段抽出法、層別抽出法などが用いられるが、古典的テスト理論に基づく研究を行う限り、サンプルの特徴によってテストの特徴が決定づけられてしまうこと、あるいは、調査結果がサンプルの質に左右されてしまうという致命的な限界が生じてしまうことは避けられない（豊田, 2002a; 渡辺, 1989）。さらに、特に組織行動測定分野においては、個人・集団・組織属性の差異が大きく、綿密なサンプリング自体が困難であることが指摘される（渡辺, 2003）。また、実際に前述の方法を用いてサンプリングを行うことは時間的・金銭的負担が非常に大きいため、より簡便な方法、例えば学生や知人など身近な人を対象に調査を行う場合が多い。本論文で問題とする、労働価値観測定尺度の開発過程における協力者の選択方法もこれに相当し、「この地域の労働者、あるいは限定される職種にしか適用できない尺度である」との批判は免れない。

項目反応理論

このように、古典的テスト理論においては、テスト（尺度）の統計量が母集団、あるいは母集団を代表すると考えられるサンプルのデータに依存するという限界がある。これに対して、項目反応理論（IRT: Item Response Theory）では、テスト（尺度）の統計量が母集団、あるいは母集団を代表すると考えられるサンプルのデータとは独立に定義される。これがIRTの最大の特徴とされる（野口, 1999）。さらに、古典的テスト理論に基づく信頼性と妥当性の検討だけでは十分に把握することが難しかった質問項目の「質」の問題にも言及できる可能性が指摘される（田中, 2004）。

IRTの基本的な考え方は、鋳物の硬さを示すモースの硬度計と同じである（芝, 1991）。標準鋳物の硬さが問題の難易度や「はい」の言いやすさに該当する。古典的テスト理論に基づいて開発されたテストでは、受験者の集団によって平均値

や分散が異なるだけでなく、異なる項目から構成される結果の比較はできない。それに対して、IRTに基づいて開発されたテスト（例えば、TOEFL）は、いつ受験したか、どこで受験したか、どの問題を解いたかに関係なく、そのスコアは比較可能になる。つまり、スコアは絶対的な英語力を表すことになる。ただし、モースの硬度計の場合は傷が「つく」か「つかない」かのいずれかであるが、テストの場合は難しい問題に正解したのに、簡単な問題を間違えることもあるため、確率モデルが導入されている（豊田，2002b）。

IRTのモデルとして心理尺度等で多く用いられるのは、2パラメタ・ロジスティック・モデルである。2パラメタ・ロジスティック・モデルでは、後述する識別力パラメタ a と困難度パラメタ b の2つのパラメタを推定することから、このように言われる。この他に、困難度パラメタ b のみを推定する1パラメタ・ロジスティック・モデル（ラッシュ・モデル）や、識別力パラメタ a と困難度パラメタ b に加えて、学力検査等の多枝選択形式のテストで生じやすいとされる当て推量パラメタ c を推定する3パラメタ・ロジスティック・モデルなどがある。

2パラメタ・ロジスティック・モデルにおいて、ある質問項目の正答確率 P は次式で表される（芝，1991；豊田，2002a）。

$$P_j(\theta) = \frac{1}{1 + \exp[-1.7a_j(\theta - b_j)]} \dots\dots (式1)$$

$P_j(\theta)$: 項目 j の正答確率
 θ : 項目 j の潜在特性尺度値
 a_j : 項目 j の識別力パラメタ
 b_j : 項目 j の困難度パラメタ

IRTでは、測定しようとしている構成概念を量的変数として捉えたとき、これを潜在特性という。式1で θ は、被験者の項目反応パターン（各質問項目ごとの回答結果）から推定される被験者の潜在特性尺度値を表す。質問項目の性質は、識別力パラメタ a と困難度パラメタ b によって表される。また、質問項目の特性は項目特性曲線によって記述される。項目特性曲線は潜在特性尺度値とその項目に正答する確率（心理テストなどでは「はい」や「あてはまる」と回答する確率）との関係を表し、潜在特性尺度値の単調増加関数になっている。これは、潜在特性尺度値が大きい被

験者の方が小さい被験者よりもその項目に正答する確率が高いことを表す。また、識別力パラメタ a は、項目特性曲線の立ち上がりの強さを表す。識別力パラメタ a が大きいほど立ち上がりは急であり、変曲点の前後で潜在特性尺度値の違い、つまり個人差を明確に識別できることになる（図1）。困難度パラメタ b は、横軸（潜在特性尺度値）上の項目特性曲線の位置を表す。識別力 b が大きいほど項目特性曲線の変曲点が右寄りであることを意味し、正答する確率が低い、すなわち難しい（「はい」や「あてはまる」と回答されにくい）問題であることを意味する（図2）。

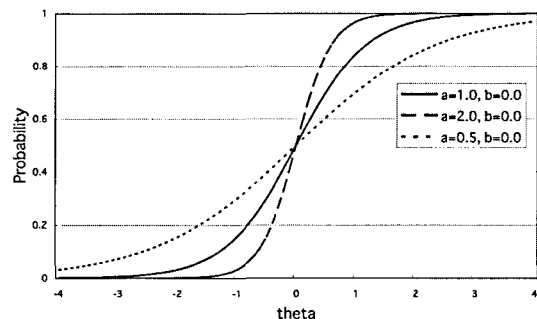


図1 2パラメタ・ロジスティック・モデルにおける識別力パラメタの影響

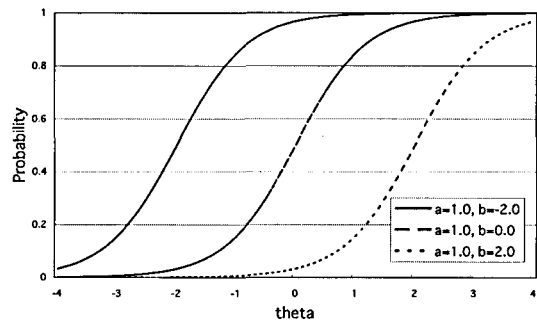


図2 2パラメタ・ロジスティック・モデルにおける困難度パラメタの影響

本研究の目的

以上のように、IRTでは、各項目の困難度や識別力がサンプルとは独立して定義される点が最大の特徴とされる（渡辺，2003）。したがって、IRTに基づいて労働価値観測定尺度を分析し、そこから得られた各パラメタが適切な値を示すのであれば、特定の地域や職種の労働者のみが本尺度の適用対象とされるのではなく、一般的な労働者を対象としたツールとしても有効であることが

示唆される。さらに、今後、本尺度を用いた応用研究を進めていく上で、質問項目の項目特性を明らかにしておくことには大きな意義があると考えられる。そこで、本研究ではIRTに基づき、労働価値観測定尺度の項目特性について分析を行い、その結果を踏まえて今後の課題についても論じた。

方法

調査方法

測定尺度 労働価値観測定尺度（江口・戸梶，2010），または労働価値観測定尺度短縮版（江口・戸梶，2009a）を用いて調査を行った。「仕事をしていく上で何を大切にしているかは人によって異なります。次の各文の内容があなたにとってどの程度重要であるか、【回答の例】のように当てはまる程度のところに○をつけてください。現在どの程度満足されているかということや、社会的に重要であると考えられている程度ではありません。現在のあなたにとっての重要度をお答えください。」と教示文を提示した。各質問項目に対して、「非常に重要である」、「重要である」、「どちらかといえば重要である」、「どちらかといえば重要でない」、「重要でない」、「まったく重要でない」の6件法で回答を求め、それぞれ6点～1点を配点した。

調査手続き プライバシー保護のため、調査は無記名で実施された。また、調査実施時に口頭、あるいは文書によって、調査への回答は強制ではないこと、個人が特定されるような形でのデータの分析は行わないこと、分析終了後には調査票は細断して破棄することなどを説明した。

対象者 2005年から2008年に大学の夜間主コースに在籍する学生および一般の社会人を対象として実施した調査から、労働価値観測定尺度短縮版に該当する21項目のデータを抽出し、1868名のデータが分析に用いられた。

分析対象者1868名（ $M = 35.6$ 歳， $SD = 12.4$ ，範囲18～74歳）のプロフィールは以下のとおりであった。性別の内訳は、男性666名（ $M = 39.9$ 歳， $SD = 13.9$ ，範囲18～71歳），女性1188名（ $M = 33.2$ 歳， $SD = 10.7$ ，範囲18～74歳），不明14名であった。学歴については、中学卒7名，高校卒389

名，専門学校卒555名，高専・短大卒178名，大学卒647名，大学院卒60名，不明32名であった。就業の経験については、現在就業中1782名，就業経験有り・現在無職49名，就業経験無し23名であり，現在就業している人の雇用形態は，正規雇用1129名，非正規雇用・フルタイム勤務41名，非正規雇用・パートタイム勤務131名，その他・不明481名であった。職位については，一般職1185名，係長・主任職クラス254名，課長職クラス202名，部長職クラス以上120名，不明21名であり，職種は技術職66名，事務職228名，管理職154名，営業職60名，販売・サービス職131名，医療職956名，作業職49名，教育職55名，自営等10名，不明73名であった。現在の会社での在職期間は $M = 10.6$ 年， $SD = 10.8$ ，通算での就業期間は $M = 13.0$ 年， $SD = 11.3$ であった。

分析方法

労働価値観の因子構造 労働価値観の因子構造については，上位因子を想定したモデルよりも，各下位尺度間に相関関係を想定したモデルの適合度が高いことが確認されている（江口・戸梶，2007a，2009a）。したがって，本研究では労働価値観測定尺度の7下位尺度ごとに分析を行うことにして，クロンバックの α 係数によって各下位尺度の内的一貫性を確認した。

各下位尺度の一次元性の確認 IRTでは，尺度が測定しようとする潜在特性が一次元であることが求められる。一次元性を確認するためには，尺度に含まれる項目のテトラコリック相関係数による項目間相関行列を因子分析し，スクリー・テストの結果，因子数が1になればよいとされる（野口，1999）。本研究では，下位尺度ごとに一次元性の確認を行った。分析には，熊谷（2009）によるEasyEstimationシリーズの一次元性確認プログラムEasyOneFactor Ver0.2.0を使用した。

項目パラメタの推定 学力検査等で起こりがちな当て推量は生起しにくいと考えられることから，モデルは2パラメタ・ロジスティック・モデルを採用し，周辺最尤推定法によって各項目の識別力パラメタ a と困難度パラメタ b を算出した。分析には，熊谷（2009）によるEasyEstimationシリーズの項目パラメタ推定プログラムEasyEstimation Ver0.4.3を使用した。また，項目

特性曲線の描画には項目特性曲線・テスト情報量表示プログラム EasyInfo Ver0.1.2を使用した。

なお、労働価値観測定尺度短縮版は、「非常に重要である」～「まったく重要でない」の6件法で回答を求めるが、本研究では「非常に重要である」～「どちらかといえば重要である」を「重要である」(=1)に、「どちらかといえば重要でない」～「まったく重要でない」を「重要でない」(=0)に二値化して分析を行った。多段階評定データについては、評定結果をそのまま用いることが望ましいとされる(大塚・小杉, 2006)。しかしながら、本研究で利用した EasyEstimation シリーズのプログラムでは、多値型モデルのパラメタ推定結果の妥当性が十分に検証されていないため(熊谷, 2009)、本研究ではデータを二値化して分析を行った。

結 果

各下位尺度の内的一貫性

クロンバックの信頼性係数は $\alpha = .809 \sim .877$ であり、十分に高い内的一貫性が確認できた(表1)。

表1 各下位尺度の信頼性係数 α

下位尺度	α
自己の成長	.840
達成感	.809
社会的評価	.863
経済的報酬	.857
社会への貢献	.842
同僚への貢献	.877
所属組織への貢献	.825

各下位尺度の一次元性の確認

下位尺度に含まれる項目のテトラコリック相関行列を因子分析し、各下位尺度の一次元性を確認した。表2に示すように、いずれの下位尺度も固

表2 各下位尺度の固有値の変化

	固有値		
	第1因子	第2因子	第3因子
自己の成長	2.550	0.301	0.149
達成感	2.423	0.317	0.261
社会的評価	2.593	0.275	0.132
経済的報酬	2.496	0.361	0.143
社会への貢献	2.498	0.292	0.209
同僚への貢献	2.666	0.178	0.156
所属組織への貢献	2.467	0.280	0.253

有値の変化から、IRTを適用する際に求められる尺度の一次元性が確認できた。

項目パラメタの推定

前述したように、2パラメタ・ロジスティック・モデルを採用し、下位尺度ごとに項目パラメタ値の推定を行った(表3～9)。

「自己の成長」尺度については、識別力パラメタは WV05が最も高く $a = 3.531$ であり、WV19が $a = 1.474$ で最も低かった。困難度パラメタは $b = -1.451 \sim -1.612$ で、いずれも負の値であった。「達成感」尺度については、識別力パラメタは $a = 1.455 \sim 1.827$ であり、3項目とも同程度であった。困難度パラメタはいずれも負の値であった ($b = -0.553 \sim -1.189$)。「社会的評価」尺度については、識別力パラメタは最も低い WV03が $a = 1.424$ で、最も高い WV10が $a = 2.947$ であった。困難度パラメタはいずれも正の値であった ($b = 0.300 \sim 0.768$)。「経済的報酬」尺度については、識別力パラメタは $a = 1.230 \sim 2.923$ であった。困難度パラメタは $b = -1.209 \sim -2.095$ といずれも負の値であった。「社会への貢献」尺度については、識別力パラメタは $a = 1.570 \sim 2.276$ であった。困難度パラメタは $b = -0.512 \sim -1.025$ とすべて負の値であった。「同僚への貢献」尺度については、識別力パラメタは $a = 2.135 \sim 2.448$ 、困難度パラメタも $b = -0.182 \sim -0.379$ と3項目とも同程度の値であった。「所属組織への貢献」尺度については、識別力パラメタは $a = 1.650 \sim 1.799$ と3項目とも同程度であった。困難度パラメタは WV06が $b = 0.384$ と正であったが、WV13と WV20は負の値であった ($b = -0.067, b = -0.409$)。

以上のように、識別力パラメタは、最も高い項目が $a = 3.531$ (WV05)、最も低い項目でも $a = 1.230$ (WV21) であり、全般的に高い値であった。また、困難度パラメタは、 $b = 0.768 \sim -2.095$ で、21項目中17項目が負の値であったのに対して、正の値を示したのは WV03, WV06, WV08, WV13の4項目だけであった。図3～9に、下位尺度ごとの項目特性曲線を示す。

表3 「自己の成長」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV05	自分自身の成長のために働くこと	3.531	-1.612
WV12	人間として成長するために働くこと	1.839	-1.716
WV19	自分の能力を開発するために働くこと	1.474	-1.451
	パラメタ平均値	2.281	-1.593

表4 「達成感」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV01	自分の持っている力を「すべて出しきつた」と思うために働くこと	1.455	-0.553
WV08	「精いっぱい働いた」という感じをもつために働くこと	1.827	-0.684
WV15	仕事において何かをやりとげたという感じをもつために働くこと	1.605	-1.189
	パラメタ平均値	1.629	-0.809

表5 「社会的評価」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV03	自分の知識や技能について、他の人々からほめられるために働くこと	1.424	0.300
WV10	多くの人に注目し、尊敬してもらうために働くこと	2.947	0.451
WV17	他の人々から注目されるために働くこと	2.375	0.768
	パラメタ平均値	2.249	0.506

表6 「経済的報酬」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV07	よい生活をするのに十分な賃金をかせぐために働くこと	2.339	-1.654
WV14	多くの収入を得るために働くこと	2.923	-1.209
WV21	人並みの生活ができるくらいの収入を得るために働くこと	1.230	-2.095
	パラメタ平均値	2.164	-1.653

表7 「社会への貢献」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV04	社会の人々を助けるために働くこと	1.647	-1.025
WV11	社会のために働くこと	2.276	-0.790
WV18	社会の幸福と平和のために働くこと	1.570	-0.512
	パラメタ平均値	1.831	-0.776

表8 「同僚への貢献」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV02	同僚の役にたつために働くこと	2.135	-0.379
WV09	同僚を援助するために働くこと	2.300	-0.240
WV16	同僚の役に立つ人間になるために働くこと	2.448	-0.182
	パラメタ平均値	2.294	-0.267

表9 「所属組織への貢献」尺度の識別力パラメタと困難度パラメタ

項目番号		識別力 (a)	困難度 (b)
WV06	所属する組織に自分を捧げるために働くこと	1.666	0.384
WV13	所属する組織のために力を尽くしていると実感するために働くこと	1.799	-0.067
WV20	仕事を通して所属する組織へ貢献するために働くこと	1.650	-0.409
	パラメタ平均値	1.705	-0.030

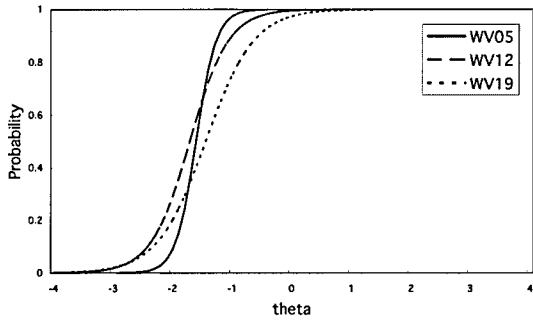


図3 「自己の成長」尺度の項目特性曲線

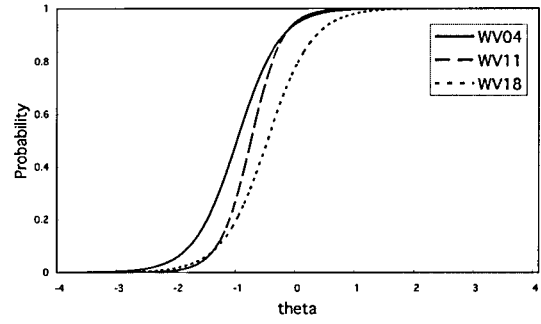


図7 「社会への貢献」尺度の項目特性曲線

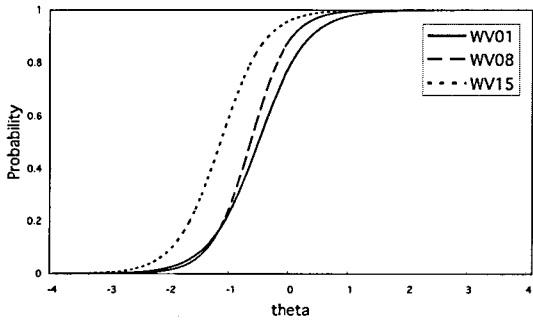


図4 「達成感」尺度の項目特性曲線

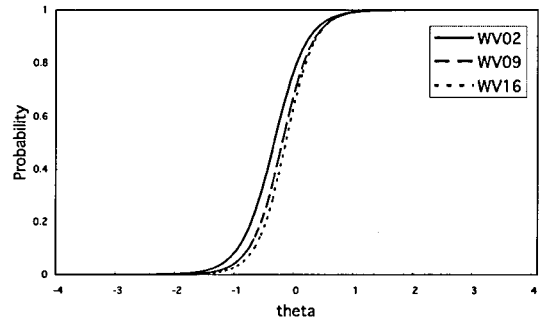


図8 「同僚への貢献」尺度の項目特性曲線

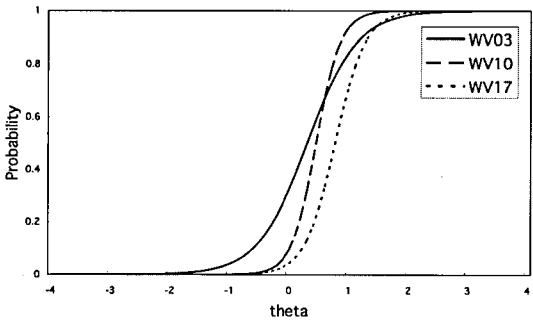


図5 「社会的評価」尺度の項目特性曲線

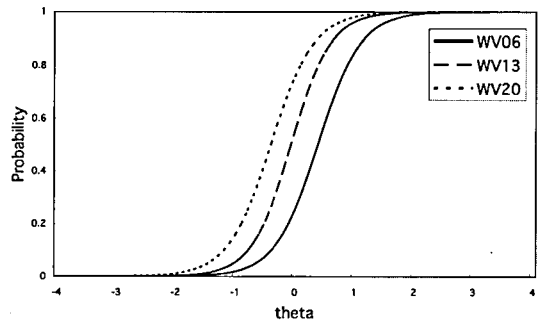


図9 「所属組織への貢献」尺度の項目特性曲線

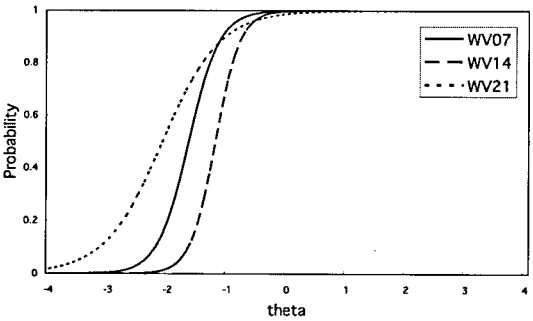


図6 「経済的報酬」尺度の項目特性曲線

考察

本研究は、労働価値観測定尺度短縮版が、特定の地域や職種の労働者のみが適用対象とされるのではなく、一般的な労働者を対象としたツールとしても有効であるかを確認するために、IRTを用いて尺度の項目特性について分析を行ったものである。

下位尺度の一次元性について

前述したように、IRTを用いる際には、すべての項目が単一の潜在特性のみを測定しているという一次元性が仮定される必要がある。表1、表2に示すように、いずれの下位尺度も高い内的一貫性と一次元性が確認され、IRTが適用可能であることが示された。

識別力パラメタについて

推定された識別力パラメタについて、どの水準で識別力が満足であると判断するかについては様々な議論がある(矢富, 1999)。例えば、これまでの研究において $a > 0.40$ (木島, 1999)、 $a > 0.50$ (Roznowski, 1989)、 $a > 0.75$ (Ironson, Smith, Brannick, Gibson, & Paul, 1989)など、いくつかの基準が提案されている。本研究では、最も高い項目(「自己の成長」尺度: WV05「自分自身の成長のために働くこと」)で $a = 3.531$ 、最も低い項目(「経済的報酬」尺度: WV21「人並みの生活ができるくらいの収入を得るために働くこと」)でも $a = 1.230$ であり、すべての項目がこれらの基準を上回った(表3~9)。特に「同僚への貢献」尺度は3項目とも $a > 2.00$ と高い値であった。したがって、本尺度の質問項目は、一般的な労働者を対象として高い識別力を有すると判断された。

困難度パラメタについて

困難度パラメタについては、 $4.00 > b > -4.00$ が適切な範囲とされる(木島, 1999; Roznowski, 1989)。本研究では、困難度は $b = -2.095 \sim 0.768$ であり、すべての項目がこの範囲に収まっていた(表3~9)。ただし、21項目のうち正の値であったのは、「社会的評価」尺度のWV03「自分の知識や技能について、他の人々からほめられるために働くこと」、「所属組織への貢献」尺度のWV06

「所属する組織に自分を捧げるために働くこと」、「達成感」尺度のWV08「『精いっぱい働いた』という感じをもつために働くこと」、「所属組織への貢献」尺度のWV13「所属する組織のために力を尽くしていると実感するために働くこと」の4項目だけであった。特に「自己の成長」尺度と「経済的報酬」尺度はすべての項目が $b < -1.00$ であった。この結果は、これらの尺度(項目)が「重要である」と回答されやすい傾向にあることを意味する。逆に「社会的評価」尺度はすべて正の値であり、「重要でない」と回答されやすい傾向にあることが示された。このように、全般的に「重要である」と回答されやすい傾向が認められたが、極端に困難度が高かったり、低かったりする項目はなく、すべての項目が適切とされる範囲に収まっていた。したがって、本尺度の質問項目は、一般的な労働者にとって適切な困難度であると判断できよう。

総括

以上の結果を総括すると、困難度パラメタは多くの項目が負の値を示したが、すべての項目の識別力パラメタ、困難度パラメタが適切とされる範囲に入っており、問題となる項目はなかった。このことから、労働価値観測定尺度短縮版は、特性尺度値が比較的低い個人を非常に敏感に識別できる尺度であると考えられる。さらに、特定の地域や職種の労働者のみが本尺度の適用対象とされるのではなく、一般的な労働者の労働価値観測定に適した尺度であることが示唆された。

ただし、前述したように、本研究では多段階評定データを0、1の二値データに変換して分析を行っており、この方法ではデータが有する情報の損失が大きい可能性が指摘される(鈴木・豊田・小杉, 2004)。今後は、多段階評定データでの検討も行う必要がある。

今後の課題

労働価値観測定尺度を用いた応用研究や現場での活用を進めていく上で、本尺度の項目特性を明らかにできたことには大きな意味がある。この成果を踏まえて、これからの課題について述べる。

第一に、項目プールの作成が挙げられる。労働

の現場での本尺度の活用方法として、例えば、労働価値観を測定することで不適応問題の理解や予防を図ることを考えたとき（江口・戸梶，2010）、同一被験者に同じ尺度を用いて、繰り返し測定することが想定される。この際、被験者が項目の内容を覚えてしまう（学習してしまう）ことによる弊害が考えられる（豊田，2002b）。それに対して、項目パラメータが既知である質問項目をあらかじめ多数用意しておく（項目プールの作成）、必要に応じて適切な項目を組み合わせる使用することができれば、同一の被験者への繰り返し測定による問題に対応できる。さらに、項目プールを作成しておくことは、コンピュータを用いた適応型テスト（CAT: Computerized Adaptive Test）への応用にもつながる。CATでは、測定対象に応じて適切、かつ必要最低限の項目を項目プールから選び出すだけでなく、被験者のレベルに応じた精度の高い測定が可能になるとされる（豊田，2002a）。また、労働現場での測定を想定すると、勤務形態などによっては集団での実施が難しい場合もあろう。そのような場合も、コンピュータを使うことで個別での測定が可能になるだけでなく、結果の集計やフィードバックも容易に行える。以上のことから、項目プールの作成とCATへの展開は、本尺度の現場での活用を考えるうえでも非常に重要なテーマであろう。

第二に、イーミック-エティック・ジレンマの解消に向けた取り組みが挙げられる。これまで、異なる集団や組織、職種に属する労働者の意識を測定し、比較するには、両者の意識を共通に反映するような一般的な項目が用いられてきた。しかしながら、それぞれの集団に特徴的な内容を含む項目で測定できれば、より現状を的確に把握することができる。ところが、従来の古典的テスト理論では、たとえ同一の概念を測定しているとしても、その理論的背景から、異なる項目を用いて測定した結果を比較することはできない（渡辺，1989）。

イーミック（emic）とエティック（etic）はアメリカの言語学者パイク（K. L. Pike）が提起した概念である（池田，2009）。イーミックとは、人間の行動をその人間が属する文化のなかにある固有の視点から総合的に研究するアプローチであり、エティックとは、人間の行動をすべての文化

を越えた普遍的な視点から研究するアプローチである。古典的テスト理論を用いて、前述のような労働者の意識を測定することを考えたとき、ある特定の集団や組織においてだけ見られる特殊性・特異性を重視すれば（イーミック）、他の集団や組織との比較が困難になる。逆に、多くの集団や組織において見られる一般性や普遍性を追求すれば（エティック）、個々の集団や組織の特徴が捉えにくくなる。古典的テスト理論に基づく研究を行う限り、このようなジレンマが避けられない。

特に、組織行動測定分野においては、個人・集団・組織属性の差異が大きいことから（渡辺，2003）、このような普遍性と特異性のバランスが常に問題になる。例えば、Cooper & Marshall（1976）によって提唱された労働ストレスに関する因果関係モデルにおける問題が有名であろう。このモデルに基づく多くの研究は、ある特定の職業に関連することが予想されるストレス項目を作成し、因子分析により抽出された因子をストレス変数とした重回帰分析でストレス反応変数との関連を検討するものであった。したがって、個々の研究対象集団においてのみモデルの妥当性が成立するという、統計学的には正確であっても、極めて比較可能性に乏しい研究ばかりが蓄積され（岩田，1997）、研究のなされていない目新しい職種を探し求める拡散的状况を引き起こしたとして批判されてきた（矢富・渡辺，1995）。しかしながら、IRTを用いることで、同じ概念を反映する異なる質問項目の比較が可能になり（鈴木・豊田・小杉，2004；渡辺，1989；矢富・渡辺，1995）、このようなジレンマを乗り越えることができよう。労働価値観概念についても、本研究の成果をベースに、組織や職種の特徴を反映するような項目を加えていくことで、イーミック-エティック・ジレンマの解消が期待される。

第三に、国際的な研究における尺度の翻訳等価性の問題への対処が挙げられる。国際的な比較研究を行う際には、原版の質問項目を翻訳した後、原版の言語に再翻訳するバック・トランスレーションが行われることが多い。これによって、翻訳時の誤訳や誤解を避けることができるが、直訳的表現が多くなり、翻訳先の言語には無い概念を無理矢理訳してしまうことで、結果的に質問の意味が分かりにくい、不適切な翻訳になってしまう

ことがある。そのため、バック・トランスレーションだけでは十分な翻訳等価性は確保できないとされる（藤本・加藤，1999；木島・高橋・野口・渡辺，1998）。この問題についても、IRTを適用することで、翻訳等価性を確認できるだけでなく、特異項目機能の問題をも解決できる可能性がある。

最後に、中長期的な課題として、尺度の改訂に関する問題がある。労働価値観は社会と個人の相互作用の中で変化していくと考えられており（林，1985），社会の中での労働の位置づけや意味合いが変化すれば，労働価値観も変化していくと推測される。したがって，将来的には労働価値観概念や因子構造の変容についても再検討する必要が指摘されており（江口・戸梶，2009a），その際には，本研究で明らかにされた項目特性を参考に改訂作業を行うことができよう。

引用文献

- Adkins, C. L., & Naumann, S. E. (2001) Situational constraints on the achievement-performance relationship: A service sector study. *Journal of Organizational Behavior*, 22, 453-465.
- Chiu, R. K., & Kosinski, F. A. Jr. (1995) Chinese cultural collectivism and work-related stress: Implications for employment counselors. *Journal of Employment Counseling*, 32, 98-110.
- Cooper, C. L., & Marshall, J. (1976) Occupational sources of stress: a review of the literature relating to coronary heart disease and mental ill health. *Journal of Occupational Psychology*, 49, 11-28.
- 江口圭一・戸梶亜紀彦（2004）ストレス研究の歴史の概観：労働ストレス研究の新しい視点を目指して 広島大学マネジメント研究, 4, 195-208.
- 江口圭一・戸梶亜紀彦（2005）労働価値観測定尺度開発のための展望 広島大学マネジメント研究, 5, 147-152.
- 江口圭一・戸梶亜紀彦（2006a）労働価値観測定尺度の開発（その1）日本グループ・ダイナミクス学会第53回大会論文集, 156-157.
- 江口圭一・戸梶亜紀彦（2006b）労働価値観測定尺度の開発（その2）：組織市民行動との関連から見た尺度の妥当性 産業・組織心理学会第22回大会発表論文集, 77-80.
- 江口圭一・戸梶亜紀彦（2006c）労働価値観とコーピングの関連についての研究 産業ストレス研究, 14 (1), 39.
- 江口圭一・戸梶亜紀彦（2007a）労働価値観測定尺度の因子的妥当性に関する検討 広島大学マネジメント研究, 7, 37-47.
- 江口圭一・戸梶亜紀彦（2007b）労働価値観測定尺度短縮版の検討 日本グループ・ダイナミクス学会第54回大会発表論文集, 82-83.
- 江口圭一・戸梶亜紀彦（2007c）労働価値観測定尺度の妥当性に関する検討：達成動機との関連 日本健康心理学会第20回記念大会発表論文集, 9.
- 江口圭一・戸梶亜紀彦（2009a）労働価値観測定尺度（短縮版）の開発 実験社会心理学研究, 49 (1), 84-92.
- 江口圭一・戸梶亜紀彦（2009b）職業生活におけるストレスと労働価値観が組織市民行動に及ぼす影響 産業・組織心理学会第25回大会発表論文集, 19-22.
- 江口圭一・戸梶亜紀彦（2009c）労働価値観の安定性に関する研究 日本健康心理学会第22回大会発表論文集, 138.
- 江口圭一・戸梶亜紀彦（2010）労働価値観測定尺度の開発 産業・組織心理学研究, 23 (2), 145-154.
- 藤本哲史・加藤尚子（1999）国際比較研究へのアプローチ：組織コミットメントの日米比較 渡辺直登・野口裕之（編著）組織心理測定論：項目反応理論のフロンティア 東京：白桃書房 Pp.259-276.
- 林 伸二（1985）仕事の価値 東京：白桃書房
- 池田俊憲（2009）被災観光客の避難状況に関する質的事例研究：観光安全システムの確立を目指して 地域政策研究, 12 (1), 131-149.
- Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989) Construction of a Job in General Scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology*, 74, 193-200.

- 岩田 昇 (1997) ストレッサー評価方法とその意義 2. 職業性ストレスモデルとストレッサー測定尺度の変遷 産業ストレス研究, 4, 30-34.
- 木島伸彦 (1999) パーソナリティの測定: TCI 気質・性格尺度 渡辺直登・野口裕之 (編著) 組織心理測定論: 項目反応理論のフロンティア 東京: 白桃書房 Pp.179-206.
- 木島伸彦・高橋弘司・野口裕之・渡辺直登 (1998) パーソナリティ尺度と組織社会化諸尺度との関連性: Temperament and Character Inventory (TCI) を用いて 経営行動科学, 12 (1), 31-48.
- 厚生労働省 (編) (2006) 平成18年版労働経済白書: 就業形態の多様化と勤労者生活 東京: 国立印刷局
- 熊谷龍一 (2009) 初学者向けの項目反応理論分析プログラム EasyEstimation シリーズの開発 日本テスト学会誌, 5, 117-128.
- 野口裕之 (1999) 項目反応理論の概要 渡辺直登・野口裕之 (編著) 組織心理測定論: 項目反応理論のフロンティア 東京: 白桃書房 Pp.3-58.
- 大塚泰正・小杉正太郎 (2006) 企業従業員を対象とした心理的ストレス反応尺度の項目反応理論を用いた検討 産業・組織心理学研究, 20 (1), 35-44.
- 大山七穂 (1990) 価値と規範 大坊郁夫・安藤清志・池田謙一 (編) 社会心理学パースペクティブ3 東京: 誠信書房 Pp.237-262.
- Roznowski, M. (1989) Examination of the measurement properties of the Job Descriptive Index with experimental items. *Journal of Applied Psychology*, 74, 805-814.
- 芝 祐順 (1991) 項目反応理論: 基礎と応用 東京: 東京大学出版会
- 祖父江孝男 (1971) 県民性: 文化人類学的考察 東京: 中央公論社
- 鈴木綾子・豊田秀樹・小杉正太郎 (2004) 項目反応モデルによるストレス反応尺度の構成とテスト特性曲線によるその深化の過程 心理学研究, 75, 389-396.
- 田中堅一郎 (2004) 日本版組織市民行動尺度の妥当性と信頼性, および項目特性についての検証 産業・組織心理学研究, 18, 15-22.
- 豊田秀樹 (2002a) 項目反応理論 (入門編): テストと測定の科学 東京: 朝倉書店
- 豊田秀樹 (編著) (2002b) 項目反応理論 (事例編): 新しい心理テストの構成法 東京: 朝倉書店
- 渡辺直登 (1989) 項目反応理論を用いた組織行動の測定: その概要と適用可能性 経営行動科学, 4 (2), 65-74.
- 渡辺直登 (2003) 項目反応理論による組織行動の測定とテストング 組織科学, 37 (1), 56-67.
- 矢富直美・渡辺直登 (1995) 項目反応理論による心理的ストレス反応尺度 (PSRS) の分析 経営行動科学, 10, 23-34.

Abstract

The Work Values Scale (short version) has been developed based on Classical Testing Theory. However, there is limitation in Classical Testing Theory, because the statistics drawn from the theory are dependent on the response of participants. On the other hand, Item Response Theory can estimate statistics independently of participants' responses. In the present study, the discrimination parameters and difficulty parameters of sub scales of the Work Values Scale (short version) were estimated by the two-parameter logistic model. The discrimination parameters ranged from 1.230 to 3.531. The difficulty parameters ranged from -2.095 to 0.768. Therefore, all the parameters showed appropriate, suggesting that the Work Values Scale (short version) is useful scale for general workers.