

5. MWS 実施のための標準化と基準表の作成

(1) 評価尺度作成に先立つ標準化の確認

(ア) 標準化の条件

個人の実施結果を評価するには何らかの評価尺度が必要である。それが明確でなければ恣意的な評価となり、公正さを欠くことになる。そのため、評価機能を持つ MWS の、今後の使用に対する便宜を向上させるために、何らかの評価尺度を備えて、個人の成績を照合できることが必要になる。

そこで、健常者に MWS を実施して収集したデータに基づいて、基準表を試作することが計画された。

ワークサンプル法に限らず、評価のための尺度を作成するための前提として、いくつかの側面における検討が必要である。

それらを列挙すれば、以下のようになる。

(a) 作業内容・実施方法について

機材の仕様とその一般的使用条件を一定にすることが重要である。同じ設備、道具類、材料、作業手順、生産性（量と質）の圧力、が与えられなければならない、そのために実施者が果たす役割と手順が明確にされるにはマニュアルが不可欠になる。機材の仕様や使用条件をいい加減にすると、重要な情報が失われたり、逆に誤った情報が混入しかねない。作業内容や実施方法の標準化は個人の実施結果の公正さや他者との比較を意味あるものにする最も基礎的な条件である。

(b) 作業結果処理・利用について

実施結果を職業評価の次元に結び付けるために、さらにいくつかの標準化が必要になる。

① 観察方法について

実施者が果たす役割として、作業課題実施中の行動観察に重点がある場合には、観察の焦点や記録がまちまちにならないように、より客観性の高い観察項目を採用したり、観察様式を準備するなどの工夫が必要となる。実際の観察時点でマニュアルを参照することは現実的ではないので、観察者としての練習や訓練が必要となる。

② 採点方法について

実施結果についての採点の実施者によって異なるならば、その評価の根拠は失われてしまう。多くの場合、作業結果の量的な側面の採点については問題がないが、質的な側面では問題が起き易い。そのためには、具体的な操作によって判断が（振って音が出ることによって緩みの証左とするなど）客観的に行えるような工夫が必要となる。それが難しい場合は、質的な判断の手がかりについて具体的にマニュアルに明記し、図解や写真等で補足することが必要である。

③ 記録方法について

結果を記録し、採点結果を個人の成績として表示し易くするためには、作業種目ごとの記録・整理用紙を準備することが必要である。個人成績をまとめて記入する様式やクライアントの感想や自己評価を残す様式

も考慮される余地がある。ケースファイルに綴じ込み、参照する際の便宜を考えることも現実的である。様式の一揃いがマニュアルに添付されていることが求められる。

④評価手順について

評価のために、基準表を備えて個人成績を照合し、相対評価を行うことが一般的であるが、障害者を対象とする評価用具では母集団の想定が難しく、工夫を必要とする点がある。いずれにしても、評価の手順がマニュアルの中で詳細に示され、必要となる基準が添付されていなければならない。

(イ) MWS における標準化の確認

MWS の標準化については、既に記述してきたように十分な検討を行っている。すなわち、機材の仕様を統一し、その一般的使用条件はマニュアルに含まれる指示書や、実施補助用の CD 等に明らかにされている。OA 作業課題では、実施結果がそのまま、パソコン内に記録として残され、実施者が記録すべき事柄についても、マニュアルや補助教材に示されている。採点方法については、得点に重み付けをしておらず、正しい作業の数量が得点とされるので、マニュアル等に明らかである。作業の正しさの判断が難しい「ナプキン折り」課題では、写真による例示によって信頼性を高める工夫を行っている。

(2) MWS の評価尺度作成の考え方

(ア) 絶対評価基準と相対評価基準

評価尺度を作成するという事は個人の成績を標本集団の成績の中に位置付け、相対化することを意味するが、作業内容の限定と標本集団の求め方により、絶対評価基準と相対評価基準に区別できる。絶対評価基準は、そこで位置づけられた個人の成績が評価目的に即して絶対的な意味を持つものである。例えば、旋盤工の求人に応募した労働者に、旋盤作業の課題を与えて、その成績を評価するために、就業中の旋盤工の集団に基づく基準を作成する場合などである。評価されるべき成績とは別な集団の準拠枠が存在することになり、クライテリオン (Criterion) と呼ばれ、成績が一定レベル以上かどうかの判断に対して利用され易い。

相対的な基準は、評価されるべき成績が、同一の方法で収集された他のデータと比べ、どのような位置に置かれるかを示す役割を持つ。多くの場合、そこから導かれる判断は蓋然的なもので、ノルム (Norm) と呼ぶのがふさわしい。

今回、作成しようとする基準は、対象とした健常者には作業課題と密接な実務・事務・OA 作業に関して豊富な経験や技能を持った人はおらず、相対的比較を目的にした基準に留まるが、異なる集団としての障害者の結果をこれに準拠させる点から、その利用にはクライテリオンの意味合いが加わるものとなる。

(イ) 偏差値基準とパーセンタイル基準

収集された一群のデータに基づいて、成績を評価する基準を作成しようとするとき、データ群の平均値と

標準偏差に基づいて作成する偏差値基準が一般的である。知能検査に代表される心理学的測定のための尺度の多くは、この方式の基準を備えている。

この基準は相対的比較のためには理想的なものであるが、ワークサンプル法に限らず、障害者のための評価を目的に、障害者のデータに基づいて作成するうえで隘路となるのが、「母集団が規定しにくい」「正規分布が想定できない」という点である。

そのため、ワークサンプル法による評価基準の開発に当たって多く採用されてきたのは、集団内における「順位」という統計量に基づいて作成するパーセンタイル基準である。

ここでは、対象集団の中の最も上位のパーセンタイル順位は99パーセンタイルであり、その成績の下に99パーセントの人々がおり、逆に、1パーセンタイルは最も下位の成績を意味することになる。

パーセンタイル基準は任意の集団を対象として、分布の形状に関係なく、簡単に作成ができる利点があるが、母集団の平均値に拠らないため、新たなデータが加わると基準全体が変動する可能性があり、安定性が低いという弱点がある。しかし、職務要件の時代的变化や地域の雇用条件の差を考えると、こうした流動的な基準をアップトゥーデイトに作り変えて利用してゆくほうが、就業支援をバックアップする評価にとって有利と考えられる。また、特定グループについてのデータがある程度集まった場合はその集団のパーセンタイル基準を考慮すべきと考えられよう。

(3) MWS パーセンタイル基準のための資料の吟味

ここでの基準表試作に使用する健常者データは、一定期間内に効果的に収集できるように計画して実施した作業に基づくものである。そこで、実施条件の違いや収集されたデータの特徴のために、基準の当てはめに制限が出ると、妥当性が低くなり、基準作成の意義が薄れるため、いくつかの点からの検討を行うこととした。

(ア) 基準作成の対象とする数値の確認

13作業課題について簡易版、訓練版の実施から得られている数値は、所要時間値と正答数または正答率である。これら2種類の数値が、基準作成の検討対象となった。今回の試作はレベルごとの数値ではなく、作業課題の全体結果を対象とした。

(イ) 数値データの拡散程度の吟味

実用的な段階評価を可能にする基準を得るには、ある程度以上（最低10-15段階）の数値の分散を必要とする。簡易版では、実施が全体に短時間であるため、時間差の広がりを示したのは、「文書入力」「検索修正」「作業日報集計」に留まった（図2-4-42、44、46）。正答率の分布においても同様に、「文書入力」「検索修正」「作業日報集計」の3課題であった（図2-4-43、45、47）。簡易版では課題数が少ない作業課題が多く、スクリーニング機能や、「訓練版」実施計画機能を果たすためには、時間基準とエラー内容の分析に拠ることが妥当であろう。

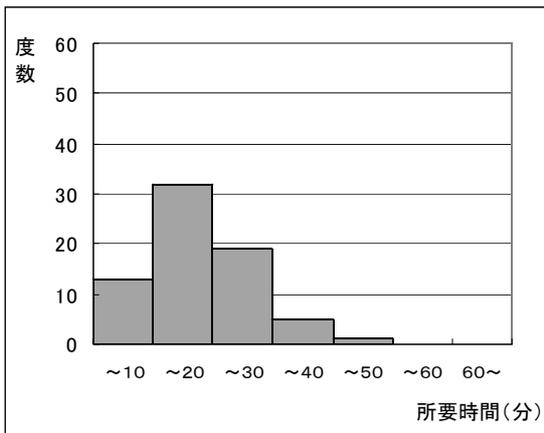


図 2-4-4 2. 「文書入力（簡易版）」所要時間

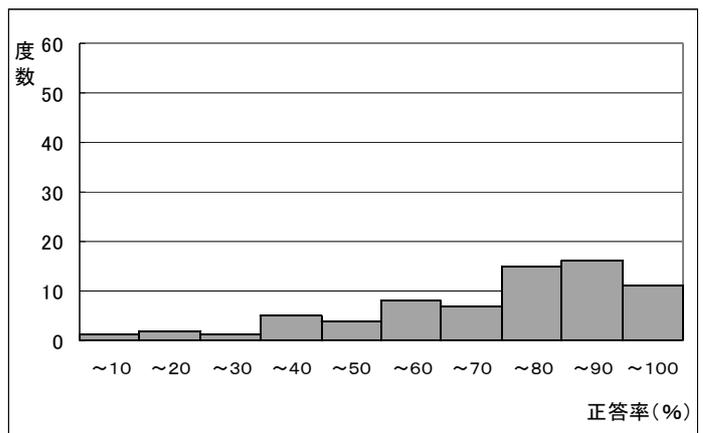


図 2-4-4 3. 「文書入力（簡易版）」正答率

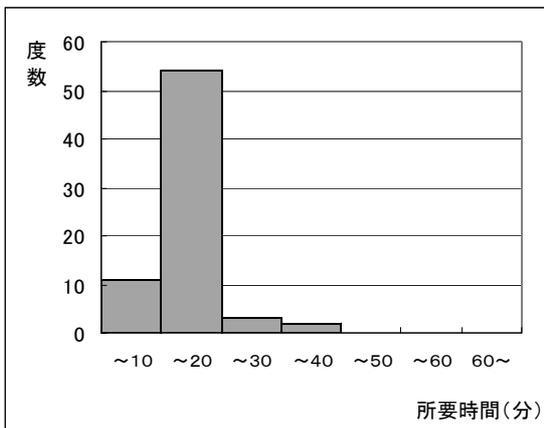


図 2-4-4 4. 「検索修正（簡易版）」所要時間

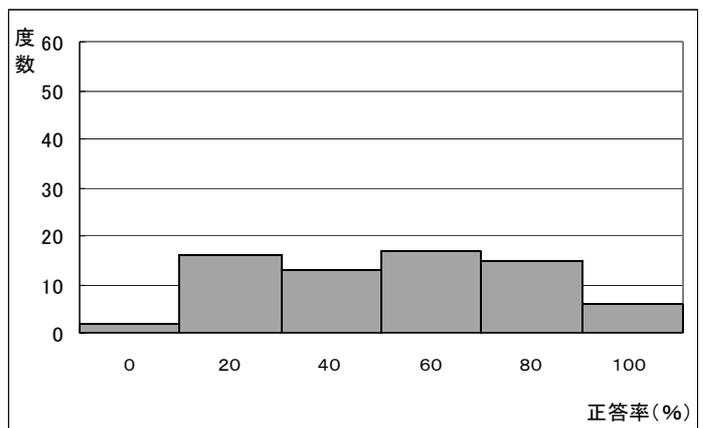


図 2-4-4 5. 「検索修正（簡易版）」正答率

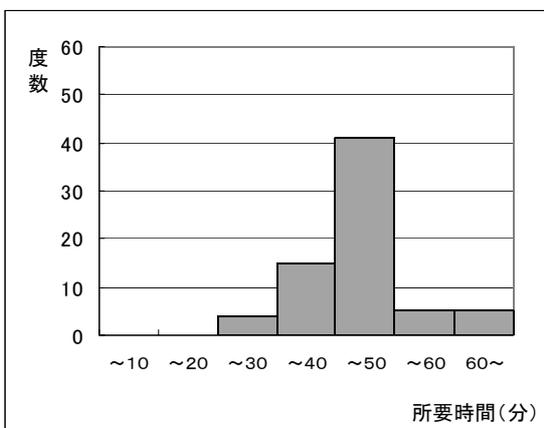


図 2-4-4 6. 「作業日報集計（簡易版）」所要時間

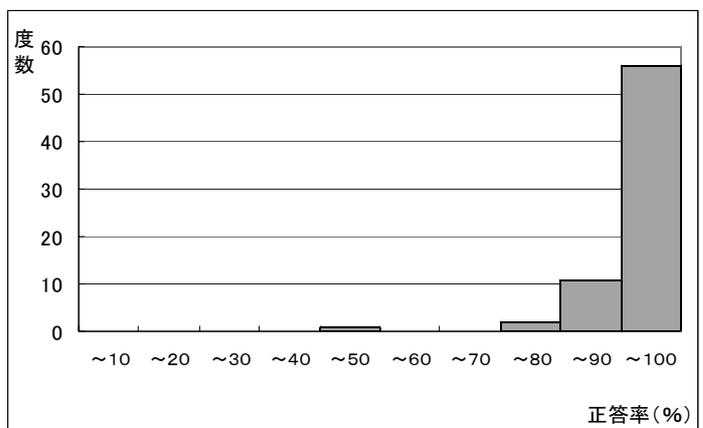


図 2-4-4 7. 「作業日報集計（簡易版）」正答率

訓練版では全ての課題の所要時間値が広く分散したが、「ファイル整理」では全員に共通実施した第1ブロックの所要時間値のみを対象とした。訓練版では「文書入力」「検索修正」(図2-4-47、48)を除く11課題で高い正答率が示され、最頻値が正答率100%に在った課題も多く、今回、それらの評価基準表を試作する意義は少ないと考えられた。

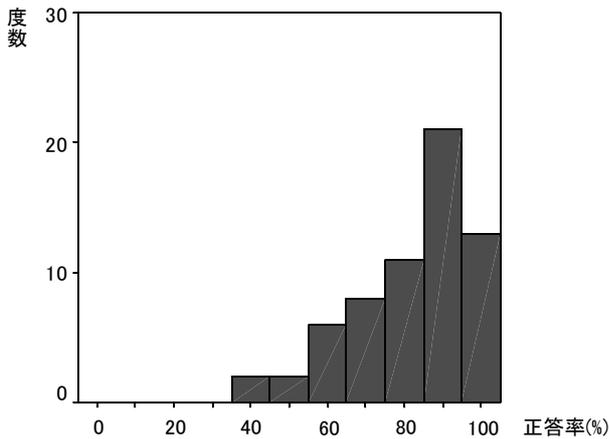


図2-4-48. 「文書入力(訓練版)」正答率

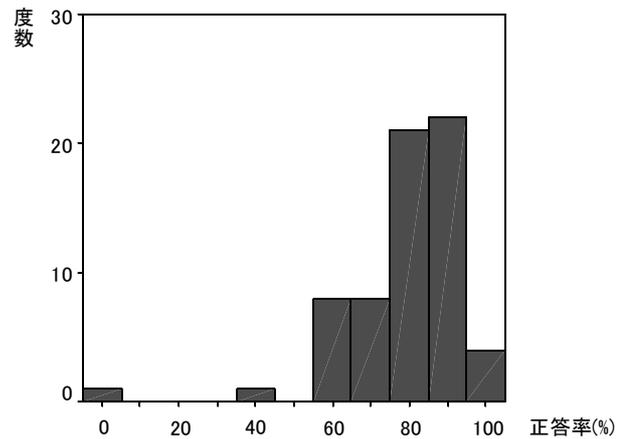


図2-4-49. 「検索修正(訓練版)」正答率

(ウ) データ内が等質かどうかの吟味

数値のバラツキの条件を充たしていても、性別の差、年代間の差が有意ならば、基準の意義は少なくなる。簡易版の時間差は「文書入力」は女性優位「作業日報集計」は男性優位だが、その差は有意ではないのでまとめて扱うこととした。簡易版の正答率には大きな差はなかった。

(4) MWSの数値に基づいた基準の試作

簡易版の「文書入力」「検索修正」では、基準表の作成課程で、同一パーセンタイル順位が多く現れたので、今回の試作から除いたため、試作基準は「作業日報集計」(表2-4-55)のみとなった。

一方、訓練版の「文書入力」「検索修正」には年代的な差が検討課題となったが、データ数が偏っているため、今回のみは

一緒に扱うこととした。将来的には「20-30代」と「40-50代」に分離して基準を作成することも考慮され

表2-4-55. 「作業日報集計(簡易版)」パーセンタイル基準表

		正答率百分位数								
		90	80	70	60	50	40	30	20	10
時間百分位数	正答率	100	99.6	98.8	98.2	97.3	95.8	93.5	90.2	90.1~
	時間(分)									
90	~32	95	87	80	70	55	50	40	30	25
80	~38	90	85	75	70	60	45	38	28	22
70	~41	88	75	70	63	57	40	35	25	20
60	~43	85	70	65	60	55	38	30	23	18
50	44	75	60	55	45	40	35	28	20	15
40	~46	65	55	50	43	35	30	25	18	10
30	47	50	45	40	35	30	25	18	15	8
20	~49	50	40	35	30	25	20	15	10	5
10	50~	40	30	25	20	15	10	8	5	3

よう。訓練版の11課題の時間基準については以下のように作成した。

表2-4-56. 訓練版11課題の所要時間基準表 (秒)

時間百分位数 作業課題名	90	80	70	60	50	40	30	20	10
数値入力	~ 184	~ 209	~ 251	~ 281	~ 309	~ 337	~ 367	~ 406	407 ~
コピー&ペースト	~ 234	~ 256	~ 278	~ 305	~ 333	~ 366	~ 413	~ 460	461 ~
ファイル整理BL	~ 239	~ 276	~ 294	~ 324	~ 347	~ 372	~ 390	~ 427	428 ~
数値チェック	~ 294	~ 347	~ 368	~ 384	~ 416	~ 448	~ 508	~ 548	549 ~
物品請求書作成	~ 2640	~ 2760	~ 2880	~ 3000	~ 3060	~ 3180	~ 3270	~ 3300	3301 ~
作業日報集計	~ 1992	~ 2132	~ 2183	~ 2337	~ 2460	~ 2572	~ 2592	~ 2873	2874 ~
ラベル作成	~ 5160	~ 5820	~ 6180	~ 6660	~ 7080	~ 7800	~ 8340	~ 9120	9121 ~
ナプキン折り	~ 6180	~ 6360	~ 6540	~ 6720	~ 6840	~ 7020	~ 7260	~ 7740	7741 ~
ピッキング	~ 854	~ 900	~ 955	~ 990	~ 1007	~ 1074	~ 1138	~ 1216	1217 ~
重さ計測	~ 327	~ 350	~ 375	~ 392	~ 413	~ 426	~ 444	~ 472	473 ~
プラグタック	~ 2299	~ 2580	~ 2745	~ 2783	~ 2927	~ 3146	~ 3305	~ 3538	3539 ~

「文書入力」「検索修正」では時間基準と正答率基準を合わせて表2-4-57、58のように試作した。

表2-4-57. 「文書入力(訓練版)」パーセンタイル基準表

		正答率百分位数									
		90	80	70	60	50	40	30	20	10	
時間百分位数	正答率	~96	~95	~92	~90	~87	~83	~75	~69	68~	
	時間(秒)										
90	~1247	95	90	85	75	55	50	40	30	25	
80	~1448	90	87	80	70	53	45	38	28	22	
70	~1586	88	85	75	63	52	40	35	25	20	
60	~1758	85	75	70	60	50	38	30	23	20	
50	~1957	75	70	55	45	40	35	28	20	15	
40	~2246	65	60	50	43	35	30	25	18	10	
30	~2726	45	43	38	35	30	25	18	15	8	
20	~3199	45	40	30	30	25	20	15	10	3	
10	3200~	35	30	25	20	15	10	8	3	2	

表2-4-58. 「検索修正(訓練版)」パーセンタイル基準表

		正答率百分位数									
		90	80	70	60	50	40	30	20	10	
時間百分位数	正答率	100	99.6	98.8	98.2	97.3	95.8	93.5	90.2	90.1~	
	時間(秒)										
90	~2326	95	87	80	75	65	50	40	25	20	
80	~2551	92	85	75	70	60	45	38	20	18	
70	~2853	90	75	70	63	57	40	35	18	17	
60	~2966	87	70	65	60	55	38	30	17	15	
50	~3084	75	60	55	45	40	35	28	15	10	
40	~3367	60	55	50	43	35	30	20	13	8	
30	~3694	40	45	40	35	30	25	18	8	5	
20	~4046	40	40	35	30	25	20	15	5	3	
10	4047~	30	30	25	20	15	10	8	3	2	

(5) 基準表を用いた評価の今後の課題

(ア) 今後の基準作成の方向

今回の基準試作はデータの不足、偏りがあったため、少数の課題のみで、パーセンタイル基準を作成した

に留まった。

今後、収集を進めて十分なデータ数を得た場合は、偏差値基準の作成を目指したい。一方、各障害者への利用データに基づく基準として、今回のようなパーセンタイル基準も用意する方がよいと考える。

(イ) 信頼性と妥当性の検証

評価機能を発揮するうえで、基本的な属性として求められるのが「信頼性」と「妥当性」である。

信頼性は、いつ、誰が、どこで実施しても目的とするものが評価できるという属性であり、評価法としての安定性、一貫性を意味する。再検査法、折半法、内的一貫性係数、代替検査法等により統計的に検証される。ワークサンプル法のような実地の作業課題を用いた評価法では、クライアントの状況変化や細かな作業条件の相違による作業成績の変動を除去することは難しい。クライアントの作業技能の向上によって成績向上が見られることも想定される。従って、再検査法による検証よりは、折半法、内的一貫性係数を利用する方が現実的である。

妥当性は、評価の対象としたものが実際に評価できているかということであり、表面的妥当性、内容的妥当性、経験的妥当性、概念的妥当性が考えられている。ワークサンプルの開発に当たっては、これらの妥当性のいくつかが検討される。しかし、ワークサンプル法は心理検査と異なり、統計的な検証を行うためのデータ収集が難しい点もあり、また、「信頼性」と「妥当性」を十分に検証してから利用を開始するという訳にはいかないため、今後、実際に利用しながら、「信頼性」と「妥当性」の検証を明確な形で行い、その有効性を向上させる努力を払うことが必要ではないかと考える。

(ウ) 臨床的使用についての情報提供

ワークサンプル法の利用においては、障害のための代替手順、方法などの使用によって、基準で想定されている事態と大きく異なることも珍しくない。そうした場合には無理に基準を当てはめることをせず、シングルケース実験法などの利用に沿って有意義な情報を集めるべきである。

そうした場合のために、実践の中での、臨床的な使用の実例も収集して、数値情報とともに提供していく必要がある。