

資 料

再検査効果についての一研究

——児童用YG性格検査の場合——

速 水 敏 彦*

問 題

教育心理学的研究では、同一人物に2回、同じ検査を実施して、その応答の変動から、心理的变化をとらえようとする研究の類が多い。そのような場合、同一の検査を繰り返すうけること自体に起因する応答の変動があるのではないかという疑いが、今まで、研究者の間にまったく無かったわけではない(石黒 1961, 岩井 1963)。しかし、その変動が具体的にどのような方向性を持ち、どの程度大きいのか、また、その種の変動が何故、生起するのか等について詳細に検討しようとした研究はほとんどない。

一方、2回の検査間に、外部からの被検者への働きかけがなく、被検者自身の心理的变化が生じていないと仮定される場合、2回の検査得点間の相関の高いことが、従来のテストによれば、信頼性が高いとされる。そこでは被検者の1回目の応答の構えと2回目のそれとが同様のものであることを前提としている。すなわち、各回の測定が同一条件下でなされ、測定の独立性が保たれることを前提としている。しかし、このような場合は、先の場合よりも、検査間隔が小さいのが通例であり、同一の検査であるという認知がより容易である。それ故、独立性は保たれず、応答の構えがかわり、規則的変動を生ずると予想される。もし、その種の変動が一定方向のものであるとするなら、信頼性係数として単に2回の得点の相関を求めただけでは、その変動を見落とし、真の値として誤って解釈する危険性が高い。

このように、同一の検査を繰り返すうけること自体に起因する応答の変動を検討する試みは、被検者の真の心理的变化をみようとして、事前テスト、事後テスト式に2回検査を実施するタイプの研究の方法論上の問題として、また、2回の単なる相関を信頼性の指標と考えてよいのかといったテスト理論の基本的問題として、きわめ

て重要である。

本研究では、検査の種類を質問紙性格検査に限定し、再検査すること自体によって応答の仕方にどのような変化がみられるのかを検討する。

Windle (1954) は再検査が実施された数多くの質問紙性格検査の結果を整理して、2回目の時点の方が適応がよいとみなされる方向に回答する場合が多いことを指摘し、それを再検査効果と呼んだ。そして、Windle (1955) はその再検査効果に関連する要因として、検査目的、検査意図についての理解、態度や雰囲気の変化等をあげている。そこで速水 (1976) は前者の検査目的、検査意図の理解の要因に注目し、矢田部、ギルフォード性格検査(以下YG性格検査と略)の下位検査である抑鬱性を測定するD尺度と一般的活動性を測定するG尺度を使って、擬装項目数と再検査による応答の仕方の変化との関連をみようとした。その結果、擬装項目数との関連は明らかにされなかったが、D尺度では、2回目の得点の方がより社会的に望ましい方向に動き、しかも、内部一致性がやや高まることが示唆された。だが、G尺度では、そのような有意な変化が見出されなかった。

そこで、まず、Windle の提唱した再検査効果という言葉で、ここでは(1) 応答が社会的に望ましい方向に変動する(再検査効果Ⅰ)(2) 尺度の内部一致性が高まる(再検査効果Ⅱ)の2つの変化をさすものであるという独自の定義を下しておく。その上で本研究では前研究をうけて、YG性格検査の全尺度をとりあげ、再検査効果の生起の様相を明らかにしたい。

ところで、再検査によって応答が社会的に望ましい方向へ変動すると考えると、大石 (1964) の研究から次のような予想がなりたつ。すなわち、おおまかにいえば、YG性格検査のプロフィール上段の情緒安定性、社会的適応性を測定するような下位尺度では得点が減少し、情緒安定性が増し、適応性が高くなる方向への変動がみられるであろう。一方、プロフィール下段の活動性、外向性等を測定する下位尺度では、得点は増大し、より活動

* 大阪教育大学

的外向的な方向への変動がみられるであろう。さらに、岩井（1966）の方向を考慮しない場合の項目の変動性に関する研究の知見に従えば、情緒安定性、神経質、内向性等といった性格の内面的な側面をあらわす項目の変動率が外面的な側面をあらわす項目の変動率よりも高いとされる。この事実も加味して考えると前述のプロフィール上段の尺度での変動の方が、下段の尺度の変動よりも顕著であろうと予想される。

次に再検査効果に関しては、検査経験に伴って、検査内容が理解され、一貫した、矛盾のない応答をしようとする構えが強くなるために、内部一致性が高まると考えられる。そして、このような構えは、特殊な質問内容に限定されることなく形成されると思われるので、再検査効果Ⅱは尺度の内容とは無関係に生起すると予想される。

さらに、上記のような2つの再検査効果をみるために、先の研究では、2回、同一の検査をして、その差を比較するにとどまったが、今回は3回実施し、再検査効果が経験回数に比例して増大するものかどうかあわせて検討したい。

方 法

被検者 大阪市立大宮小学校 5年生5クラス 男子95名、女子85名、なお、この小学校では、ここ数年間、YG性格検査を実施しておらず、被検者は未経験であると考えられた。

手続 12の下位尺度 すなわちD（抑鬱性）尺度、C（回帰性傾向）尺度、I（劣等感）尺度、N（神経質）尺度、O（主観的）尺度、Co（非協調的）尺度、Ag（攻撃的）尺度、R（のんきさ）尺度、G（一般的活動性）尺度、T（思考的外向）尺度、A（支配的）尺度、S（社会的外向）尺度を含むYG性格検査（小学2年—6年用）を1週間間隔で3回、同一人物が検査者となり教室で実施した。被検者の応答の速さは3回ともなるべく等しくなるように検査者は時間に注意して項目を読んだ。およそ必要な時間は30分であった。また、実施期日は、3回の検査中に学校側で、非常に重要な試験や行事のない昭和50年10月中旬が選ばれた。

スコアリングの方法はYG性格検査作製者の指示の通りである。各下位尺度得点は理論的には0～8に分布する。そして上の（ ）内に示した性格傾向の強いほど高い得点となる。

結 果

(1) 再検査効果Ⅰ：社会的に望ましい方向への変動

TABLE 1-a, bは各回毎の12の下位尺度得点の平均、および標準偏差であり、TABLE 1-a', b' は平均値の差の検定結果を示したものである。

TABLE 1-a Means and SDs of 12 sub-scales
for each test of 3 times (Boys)

Test M,SD Scale	1 st		2 nd		3 rd	
	M	SD	M	SD	M	SD
D	3.063	2.112	2.747	2.384	2.653	2.578
C	3.347	1.953	3.042	2.112	2.979	2.271
I	2.632	2.052	2.337	2.269	2.116	2.330
N	3.716	2.223	3.253	2.388	3.126	2.398
O	3.074	1.814	2.842	2.109	2.642	2.176
Co	2.979	2.031	3.432	2.396	3.474	2.388
Ag	5.021	1.673	4.642	1.963	4.579	1.987
R	5.200	1.850	4.853	2.067	4.705	1.999
G	3.305	1.931	3.347	2.170	3.347	2.131
T	3.368	2.157	3.263	2.263	3.232	2.232
A	4.063	1.863	4.242	1.907	4.474	2.107
S	5.221	2.027	5.337	2.106	5.579	1.998

TABLE 1-b Means and SDs of 12 sub-scales
for each test of 3 times (Girls)

Test M,SD Scale	1 st		2 nd		3 rd	
	M	SD	M	SD	M	SD
D	3.588	2.261	3.059	2.676	3.176	2.697
C	3.553	1.913	3.188	2.172	3.294	2.310
I	3.565	2.209	3.294	2.415	3.071	2.325
N	3.741	2.036	3.365	2.584	3.353	2.565
O	3.588	2.077	3.035	2.375	3.047	2.356
Co	2.812	2.059	3.000	2.346	3.071	2.505
Ag	4.800	1.651	4.388	1.923	4.318	2.082
R	4.918	1.880	4.635	2.069	4.718	2.278
G	3.106	1.980	3.141	2.143	3.294	2.179
T	3.294	1.951	3.235	2.016	3.153	2.150
A	4.000	1.891	4.094	2.079	4.118	2.089
S	5.432	1.805	5.565	2.117	5.506	2.167

まず、男子の結果についてみれば、1回目に比べて2回目では12尺度のうち半数の6尺度の得点が $P < .05$ で有意に一定方向へ変動している。さらに、1回目と3回目の比較では、 $P < .01$ で有意な変動がみられる尺度が6

TABLE 1-a' Result of testing the significance of differences between means (Boys)

Test Scale	1—2	2—3	1—3
D	N. S	N. S	*
C	*	N. S	N. S
I	*	N. S	**
N	*	N. S	**
O	N. S	N. S	**
Co	*	N. S	**
Ag	*	N. S	*
R	*	N. S	**
G	N. S	N. S	N. S
T	N. S	N. S	N. S
A	N. S	N. S	**
S	N. S	N. S	*

TABLE 1-b' Result of testing the significance of differences between means (Girls)

Test Scale	1—2	2—3	1—3
D	**	N. S	*
C	N. S	N. S	N. S
I	N. S	N. S	**
N	N. S	N. S	N. S
O	**	N. S	**
Co	N. S	N. S	N. S
Ag	**	N. S	**
R	N. S	N. S	N. S
G	N. S	N. S	N. S
T	N. S	N. S	N. S
A	N. S	N. S	N. S
S	N. S	N. S	N. S

つ、 $P < .05$ で有意な変動がみられる尺度が3つもみられる。

予想されたように、変動が大きい尺度は、I 尺度、N 尺度といった情緒安定性をみる尺度 Co や Ag 尺度のような社会的適応性をみる尺度である。G 尺度、T 尺度、A 尺度、S 尺度等の一般的に性格の外面的な側面をみる尺度では有意な変動が認められないが、例外的に R 尺度では有意である。また、その変動の方向であるが、D, C, I, N, O 尺度等は予想されたように得点が減少する方向、つまり、情緒安定性、社会的適応性が増大する方向への動きがみられる。だが、Co 尺度得点だけは増加しており、予想に反して、非協調的な方向へ変動

している。Ag 尺度、R 尺度については、社会的に望ましい方向という場合、個人差が大きく、変動の予想はむずかしいように思われるが、結果としては、いずれも得点が減少する方向、すなわち、攻撃的でなく、のんきでない方向への有意な変動が認められる。G, A, S 尺度等は変動そのものが小さく有意ではないが、予想どおり、正の方向への変動が生じている。さらに、各回の得点を比べると2回と3回のそれはかなり類似していることが明白で、2回と3回の間に有意差のみられた尺度はない。再検査効果は、回数に比例して同じ割合で強まるものではないことが理解される。

次に女子の場合であるが、有意な変動がみられる尺度の数は、男子に比べてやや少ない。もっとも $.05 < P < .10$ の水準まで許して有意差をみると、1回目と2回目の比較では、C 尺度、N 尺度、R 尺度に差がみられることになる。また、1回目と3回目の比較でも N 尺度に差がみられることになる。一方、変動の大きい尺度、あるいは変動の方向等は男子の結果と類似している。すなわち、D 尺度、I 尺度、O 尺度、Ag 尺度等ではかなり大きな得点の減少がみられるが、G 尺度、T 尺度、A 尺度、S 尺度等では差得点が小さく有意な変動はない。また、2回目と3回目の変動幅は1回目と2回目の変動幅に比べて、かなり小さなものであることも男子の結果と同様である。

さらに、尺度得点はもともと個々の項目得点から構成されているものであるが、同じ尺度に含まれる8つの項目得点は、それぞれ、尺度得点の変動と同じような変動をしているものかについても検討され、尺度得点の変動

TABLE 2 Internal consistency (α coefficients)

Test Scales	Boys			Girls		
	1 st	2 nd	3 rd	1 st	2 nd	3 rd
D	.605	.695	.737	.617	.749	.752
C	.567	.648	.688	.548	.642	.682
I	.615	.706	.736	.633	.688	.682
N	.633	.690	.694	.566	.735	.731
O	.473	.618	.651	.585	.699	.690
Co	.570	.673	.679	.646	.700	.733
Ag	.654	.506	.545	.383	.513	.596
R	.643	.594	.562	.505	.589	.651
G	.525	.642	.640	.583	.641	.645
T	.558	.645	.622	.511	.568	.626
A	.535	.565	.624	.541	.622	.593
S	.658	.609	.590	.567	.668	.667

が、2, 3の特殊な項目の変動に帰因するものでないことが確認された。

(2) 再検査効果Ⅱ：内部一致性の変化

各回の12尺度の内部一致性を α 係数で表示したものがTABLE 2である。

男子の結果では、Ag尺度、R尺度、S尺度の3尺度以外の8尺度では、1回目比べて2回目、3回目の α 係数の方が明らかに高まっている。

女子の結果は、1つの例外的な尺度もなく全尺度とも1回目よりも2回目、3回目の検査時での内部一致性の方が高くなっている。

男子と女子の α 係数を比較してみると、1回目では12尺度中6尺度で男子の値が高く、のこりの6尺度で女子の値の方が高い。続く2回目では女子の方が高い尺度が7つ、3回目では9つみられる。この事実から、内部一致性の高まりは、変動の大きさとは逆に、女子の方が男子よりも著しいといえよう。

また、これは男女ともに共通にいえることであるが、再検査効果と同様、再検査効果Ⅱでも、2回目と3回目の差ははっきりせず再検査効果は直線的に増大するものではないことが示唆された。

討 論

1. 再検査効果ⅠおよびⅡについて

検査経験に伴う応答の社会的に望ましい方向への変動は、尺度内容により異なるものか否かが本研究の1つの検討目的とされた。そこで我国の代表的な質問紙性格検査であるYG性格検査を用いて検討した結果、D尺度、C尺度、I尺度、N尺度、O尺度といった性格の内面を測定する尺度の方が、性格の外面を測定する尺度よりも、その方向への変動が一般に大きいことが明らかにされた。だが、性格の外面を測定する尺度でも、Ag尺度は男女ともかなり明瞭な変動が認められた。また、これらの尺度が全部社会的に望ましい方向へ変動しているわけではない。ただし、ここで社会的に望ましい方向というのは、一般的な判断規準によるので、児童自身の資料をもとにしていないことは注意しておかねばならない。Co尺度では予想に反して、2回目の方が得点が増加し、非協調的な方向への変動がみられた。非協調的というのは一般的に社会的に望ましくない方向といえよう。何故、予想に反してこのような例外的な変動が生じたのかについては今のところ推測できないが今後、検討する必要がある。

しかし、とにかく本研究の結果、今までほとんど注目

されなかった再検査により、一定の方向への応答の変動があるという事実、つまり各回の測定の独自性が疑わしいという事実、さらにその大きさは、尺度内容により異なるという事実は、先の速水(1976)の研究結果に比べ、より一層確証されたといえよう。なぜなら、第一に、今回は検査を3回実施し、1回目と2回目、1回目と3回目の間の変動が同一方向であったことから、検査を2回だけ実施した先の場合よりも、その変動が、検査を再びうけたこと自体による効果であるという可能性が高く、単に偶然的な条件によるものであるという可能性が低いことがあげられる。第二には、先の研究はYG性格検査の2つの下位検査だけをとりあげて検討したが、今回は全部の下位尺度について吟味したことである。

ところで、再検査効果Ⅰの変動がなぜ生じるのかについては、本研究でも明らかでない。前研究の続きとして、検査経験に伴う検査意図の理解が主要な要因であろうと最初、考えてきた。しかし、YG性格検査のように、単一尺度でなく、多くの尺度項目がばらばらに配列されているような場合でも、検査意図が理解しがたいにもかかわらず、有意な変動が認められたことから、その要因のみが決定的なものであるという考えにあまり固執するのは妥当でないように思われる。Windle(1955)があげたように、その検査事態の雰囲気の変化が鋭敏に反映されたと考えられる。すなわち、1回目の検査事態は被検者にとってなじみが薄く、緊張や不安の高まる事態であると推測されるが、その後、検査事態に慣れることによって、その種の緊張や不安が減少すると考えるならば、性格の内面を測定する尺度得点の大きな変動は、そのまま検査事態の雰囲気の変化を反映したものだといえる。

再検査効果という言葉を使うと、いかにも最初の検査時での測定値が真の値で、2回目3回目の測定値は、真の値に再検査効果の値が加わったもののように見がちであるが、別の考え方も上の例のように可能である。つまり、検査を最初にうける時こそ、特別の緊張が伴ない、そのための効果が真の値にプラスされるが、2回目、3回目には検査をうけることに慣れ、自分自身をより正確に客観視して応答できるようになり、安定した真の値に近づくとも考えられる。どちらの測定値の方が妥当性が高いかということも、今後、明らかにせねばならない。

次に再検査効果Ⅱに関してであるが、この効果は2, 3の例外的な尺度を除き、はっきり見られた。また、この再検査効果Ⅱの大きさは再検査効果Ⅰと異なり、尺度内容とはほとんど無関係であった。 α 係数の性質からい

TABLE 3 Eigen values of five factors

Factor	Boys					Girls				
	I	II	III	IV	V	I	II	III	IV	V
1 st	11.95	5.71	3.70	2.80	2.35	10.24	6.40	4.37	3.34	2.85
2 nd	16.18	6.22	4.17	3.18	2.73	16.28	7.38	4.72	3.53	3.07
3 rd	18.86	6.11	3.85	3.41	2.83	16.58	7.79	4.45	4.18	3.18

って多くの場合、尺度得点の分散の大きさは α 係数の高まりに反映するが、ここでも、TABLE 1-a, bの標準偏差の大きさと照会してみればそのことが明らかである。

すなわち、検査経験に伴ない個人の尺度得点は高い方向と低い方向に両極化する傾向があり、そのことが α 係数を高めているのである。ただし再検査効果Iとの関連で考えると両極化する際に半々に分かれずにやや社会的に望ましい方向に偏って変動すると考えられる。

また、再検査効果IIを規定する主な要因に関しても、再検査効果Iのそれと同様、今後検討していかなければならないが、再検査効果Iの場合、考えられた検査事態の雰囲気の変化のような要因は、再検査効果IIの場合は考えられないように思われる。

2. 項目間関連の変化

本研究では、再検査効果IおよびIIについて、検査作製者の指示したとおりに、12の下位尺度ごと個々独立に検討してきたが、続ら(1970)は、被検者の応答を因子分析したところ、12の特性に対応する因子が抽出されなかったとしている。検査作製者の意図した検査内容の構造と実際に測定された結果としての検査内容の構造の間にはずれがあるわけである。そこでここでも、被検者の応答から項目間の関連を一応吟味しておく必要がある。そして特に、再検査効果IやIIの影響、あるいは別の要因によって、検査経験に伴ない、項目間関連がどのように変化するのかについて注目しておくべきである。

各回ごとに96項目の応答結果を主因子解により因子分析した、男女とも5因子抽出したが固有値の大きさはTABLE 3にみるとおりである。この結果で明らかなことは、第I因子の固有値が、男子の場合も、女子の場合も1回目比べて2回目でかなり急激に大きくなっていることである。この第I因子負荷量をここでは呈示するゆとりがないが、D, C, I, N, O, Coといった尺度項目では高い正の値が、Ag, R, G, T, A, Sといった尺度項目では低い負の値がほとんど例外なく認められた*。つまり、再検査効果が比較的顕著であった尺度項目群は1

つの因子として抽出されるわけであり、その項目間の関連性が検査経験に伴ないより密になることを意味している。同一下位尺度項目間の関連性が強まることで再検査効果IIであったが、他の下位尺度項目間との関連性も変化するわけであり、これはまた別の種類の再検査効果として、今後検討を要しよう。他の4因子については、ここで特筆すべき顕著な変化はみられなかった。

要 約

YG性格検査を同一人物に1週間間隔で3回実施し、再検査を受けること自体によって生ずると考えられる応答の変動—再検査効果I（社会的に望ましい方向への変動）—と再検査効果II（内部一致性が高まる方向への変動）の生起の様相を明らかにしようとした。

再検査効果の大きさは尺度内容により異なることが指摘された、情緒安定性や社会的適応性を測定する尺度では顕著であった。

再検査効果Iの内部一致性の増大は、尺度内容とは無関連にかなり明瞭に認められた。

討論では、再検査効果がどのような要因で生起するのかといった問題や、2つの再検査効果とは別の項目間関連の変化という意味での再検査効果の問題についてもふれた。

文 献

- 速水敏彦 1976 質問紙性格検査の再検査効果 教心研 24, 54—61
- 石黒彰二 1961 中学生の学習興味—縦断的研究— 教心研 9, 160—170
- 岩井勇児 1963 性格検査項目の反応の変動性に関する研究I 教心研 11, 203—209
- 岩井勇児 1966 性格検査項目の反応の変動性に関する研究II 教心研 14, 15—24
- 大石明子 1964 個人特性評価における社会的要求性 心研 34, 29
- 続有恒・織田揮準・鈴木真雄 1970 質問形式による性格診断の方法論的吟味—YG性格検査の場合— 教心研 18, 13—47
- Windle, C. 1954 Test-retest effect on personality questionnaire. *Educ. psychol.*, 14, 617—633
- Windle, C. 1955 Further studies of test-retest effect on personality questionnaires. *Educ. psychol. Measmt.*, 15, 246—253 (1976年5月10日受稿)

* この結果を御希望の方は著者まで御連絡下さい。