

創造性検査尺度の構成と吟味

住 田 幸 次 郎

1. 本研究の要約¹⁾

我国にはテスト作成に関する情報の公開された創造性検査はみあたらない。われわれは **TC-IS** (創造性テスト第1尺度) を構成し、それについての基礎資料を提供してみたい。

この検査は言語テスト材料を用いて知的創造性を測ろうとするものであり、測定内容は芸術的創造性とは区別される。5つの下位検査にわかれている。各個人の創造力は下位検査別偏差値の総和としての創造点で示される。また、必要に応じて流暢性、柔軟性、独創性などのスコアを得ることができる。

京大男子学生により標準化された本検査には次のような傾向がみられた。

- (イ) 信頼性(折半法)は.75である。
- (ロ) 各項目は知能テストの結果とはちがった知的能力を測っている。
- (ハ) 知的行動傾向の自己評定を利用した論理的妥当性、専門課程への進級等実績との関連を調べた実際の妥当性、ともにかなり高い。
- (ニ) 性差、文科・理科系専攻者による成績の差はともにみられない。
- (ホ) 高校生の能力弁別にも大体において適用可能と考えられる。
- (ヘ) 性格、興味等に関する測度との関係についてはこの論文では省略する。

この検査は今後の研究の踏み台として使用し、創造性の研究やテスト自体の改良に努力したい。

2. 説 明

テストにより各個人の創造力を正しく測定評価することができるならば、創造性研究においても、また、個人能力の推定や将来性の予測においても、きわめて有益であろう。

しかし、我国における創造性テスト研究の現況はレベル的にはまだ乳児期とも云うべき状態である。

昭和40年10月現在において学術論文に創造性尺度の構成に関する詳しい研究結果や報告が公表されたことは皆無である。この種の報告としては僅かに日本心理学会における増田・広井の発表があげられる²⁾が、これも作成途次の経過報告と考えられ、完結したテストに関する情報は今迄なかったと云ってよい。

1) 本研究は筆者の京大大学院博士課程資格論文の一部に学術振興会奨励研究生としての研究を加え、まとめたものである。

2) 増田幸一 ○広井甫 創造性テスト作成の試み (I) 日本教育心理学会発表 1964 (II) 日本教育心理学会発表 1965

こうした現状は、検査を手段として創造性研究を試みようとする人々のためにきわめて不便であるばかりでなく、学界の創造性研究自体に対する評価にも関係すると考えられよう。

われわれがテストにより創造性研究にとり組もうとする際に、まず第一に驚ろくのはテストに関する信頼すべき情報がゼロに近いことである。必らずしも創造性測定に支持的、有効的な情報のみについてでなく、否定的、失敗例的な情報についても同様である。

われわれの研究の意義は、完成したテストを測度に用いて創造性に関する諸種の研究を進展させることにあるのは勿論であるが、直接には各種の基礎情報の提供という目的をもつ。われわれ自身、われわれの構成した尺度は不完全なものと考えている。その上、測定内容の純化を計るためとは言え、ちがった角度からみれば不自然に思える測定内容の制限や思い切った採点方法の固定化をおこなっている。しかし、こうした制限範囲内で資料を忠実に提供しえたことは満足できるように思える。以下、資料とその説明を簡単におこなってみたい。

3. テストの構成

名称 便宜上、本検査を TC-1S³⁾ (創造性テスト第1尺度) と名づける。

創造性を測るために一種または数種の創造性検査を任意にとりだして実施する、というのが従来のやり方であった。

しかし、知能テストのように、数個の下位検査をまとめた形のものとして用いたり、その結果を公表したりすることはほとんどおこなわれなかった。それは、次のような理由によるのではなからうか。

(1) 創造性の概念は甚はだあいまいであって、研究者により定義や考えかたが異なっている。誰をも納得させるような概念の設定が困難である。

(2) 参考となる基礎データがなく、基本的な研究から始めなくてはならない。創造性に関する研究は過去にも多いが、テストそのものについての分析は、特に我国においてはおろそかにされている。参考にするにしても外国のデータをそのまま用いることが多い。

(3) 創造性テストは採点の客観化、単純化がきわめて困難である。客観テスト形式、マルチョイ形式の利用できる場合は少ない。

(4) 標準化がやりにくい。適切なサンプルが得にくい。

(5) その他

検査作成に先だち、われわれは以上の問題について検討した。その結果、要約して次のような方針、手続きを踏まえることにした。

(1) 過去において用いられてきた、またわれわれが予備テストにおいて検討の対象とした多くの創造性検査を考えてみれば、知的(科学的)創造性、および芸術的創造性は、別のものとして考えるのが望ましいように思える。

3) Test of Creativity 1st Scale.

両能力は過去の多くの研究から推測されるように一種の能力として混同し用いることは困難であり、これを混同する場合には有効なテストを作りあげることができないように思える。われわれは、今回は芸術的創造力よりも一般能力的要素が濃い、また現代社会に於ける用途も広いと考える知的創造性を問題にすることに決めた。

また、形式上からみても、言語性テストと非言語性テストを混ぜて使用するよりも、今回は言語性のみに材料を限り、測定内容の明確化を意図した。

(2)は、われわれの手で予備検査、関連研究、標準化後の吟味を詳細におこなうことにより解決に近づきたい。

(3) 既に、予備検査において採点の客観化の困難さ、採点の煩雑さが伺われた。そのため、このテストでは測定単位（粗点）を正解反応の内容カテゴリーの数⁴⁾と、その非凡反応得点⁵⁾の2種類以外のものは使用しないように一応限定した。

この方法を用いれば、僅かにニュアンスの異なる反応が、場合により同一カテゴリーに分類されてしまう懸念はあるが、採点者の解釈の相違に基づく判定誤差は殆んど防止できることが証明された。

(4) 標準化集団決定の方法として、(i)対象集団（たとえば大学生）から広範囲に、偏らぬように標本抽出をおこなう (ii)対象群をできるだけ純粋なもの⁶⁾に求め、これより作成された基準を以て他の群にも適用する。の2つの場合が考えられる。一般には前者がよく用いられるが、われわれの場合には限られた数の標本下においてはどのような被験者群がどんな割合いで混入されたか、すればよいか、といった根拠が確立しにくく、標準の恒常性を保ちにくい。適性検査の標準設定に使用されることの多い後者の方法が、信頼性妥当性ともに高い尺度を得ることが可能と考えた。

集団は一定水準の知的創造力を保有すると考えるものを対象にしなくてはならない。したがって、相当の知能の高さと、ある程度の年齢を要求される。また、その集団について将来ある程度の業績をあげることが予想できることが望ましい。また、上述の条件の年次変動が小であるほどよい。これらの条件を検討した結果、京大の教養課程在学生（主として一回生）を使用することにした。念のため、標準化とはほぼ同じ状態で各学年50名ずつに創造テストを実施し、各年度間の変動が小であることをたしかめた。女子は数も少なく、創造性に関する情報も不明なので標準作成からは除外することにした。

このような経過を経て、50種余りの問題中より下位検査の選択、標準化がおこなわれた。

4) 同一カテゴリーに属する反応をいくらか多く記入しても、得点は1点と考える。たとえば、鉛筆の用途について絵を描く、字を書く、塗る、といったものは同一カテゴリーに属し、穴あけに用いるのはこれとは別カテゴリーとして扱おう。

5) カテゴリー別の頻数分布を作成し、反応の多い方からカテゴリー頻数を1位、2位、3位……と総頻数の2/3をこえるまで累積していく。2/3を累積量がこえたとき、その次位以下の反応を非凡反応とする。

6) 内的に homogeneous であるように。

(2.1) 下位テストの選択

予備検査と因子分析を繰返した結果、5つの下位検査が選ばれた⁷⁾。説明時間を除き、正味検査時間は32分。全体を1時間以内に完了するよう計画したため、当初の計画より下位検査数は制限された。

(3.2) 測定内容

粗点は各下位検査ごとに計算された偏差値に換算され、下位検査別の能力程度をあらわす。各個人の能力の代表値は5つの下位検査の偏差値合計(平均250点)で示し、これを創造点と呼ぶ。以後の研究はすべて粗点・偏差値・創造点のいずれかを単位としたものである⁸⁾。

下位検査別の時間、問題数などにつき Table 1 に示した。

7) 1. 同音異義連想(同音)

日本語には読み方が同じでも異なった意味をあらわす単語が多い。この検査では平仮名で書かれた数個の文字から、できるだけいろいろな意味を連想させる。たとえば、こいという言葉から魚のこい(鯉)、ラブ(恋)、わざとする(故意)……といった言葉を書かせる。

漢字の知識量のテストでないから解答は必ずしも漢字で答えなくてもよく、目的物を直接に連想させることばとか、云いかえた記述を記入しておけばよい。因みに、この検査とある漢字知識量が問題となるテストとの相関は0.15にすぎない。採点は25項目を通じての正答カテゴリー数。言語流暢性を測定するテスト。

2. 結果テスト(結果)

現実にはおこりえないようなことがらの生起を想定し、その結果生じることをいろいろと空想させる。たとえばもし、世の中の法律が全廃されたら……どうなるかを考えて書かせる。外国の例ではこのテストは遠隔連合の指標としても用いられているが、われわれの場合は採点法が異なるのでそのためには用いない。正答頻数プラス独創点で粗点とする。観念の流暢性を測ることをめざす、独創性の指標としても部分的に使用する。

3. 用途テスト(用途)

このテストは、身のまわりにある身近かな物について、その本来の用途以外の使いみちをいろいろあげさせることにより、柔軟性を測定しようとするものである。たとえば新聞紙は読む目的以外にものを包む、虫をたたき落す、火種として使用することができる。

検査2、同様に正答頻数プラス独創点で粗点とするが、柔軟性の最も安定した指標として使用する。

4. 問題点の発見(問題)

日常ありふれたものの欠点をあげさせ、改良点を書かせる。本検査の1. 2. 3. はともに広い可能性範囲の中から適当なものを集めさせるといった形式に反し、この検査は対象を限り、その中から課題をさがし、対象物の機能が十分発揮される上での障害となっているものをさがし出す。他の検査よりも各個人の経験を、また問題に対する平素の感受的態度を活かすことができ、年齢による衰退も小さいと考えられる。

粗点は頻数(カテゴリー数)を用い、課題に対する感受性、柔軟性を測る。

たとえば自転車の改良すべき点をあげよ、といった問題を使う。

5. 共通点の発見(共通)

2種のものを対にして呈示し、その共通点をいくつもあげさせる。採点の際には、同じような理由を何度も云いかえてのべたときにはそれらを通じて1点しか与えない。たとえば、東京と大阪の共通点をあげさせたと仮定すると、都市である、人口が多い、といった記述は同じことの言い換え、と考える。また、東京は東、大阪は西にある。といった表現の際には点を与えない。

例 リンゴとミカンの共通点をあげよ。

正答カテゴリー数で粗点とする。柔軟性の測定。

- 8) 因子構成上から云えば検査1. 2. で総合的な流暢性の、検査3. 4. 5. で柔軟性の、検査2. 3. の非凡反応得点で独創性の測度とすることも可能と考えられる。この点の分析や検討が不十分なのと混乱をさけるため本文ではとりあげない。

(3.3) 標準化の実施

昭和37年9月～10月に京大教養部より各学部混成の講座7クラスを選び、受講生に標準化を実施した。抽出は各講座にばらつく

くようにして、特定クラスに偏らないようにし、男子290名を対象に標準化を実施した。第2, 第3検査において平凡反応カテゴリー, 非凡反応カテゴリーの区別を定める必要があるので、やや実施時期が他と離れた40名を除き、250名の結果より標準を定めた。

Table 1 下位検査の実施時間と問題数

下位検査(略号)	時間	問題数	採点法	仮想因子
1. 同音異義連想(同音)	7分	25	f	流暢
2. 結果テスト(結果)	7分	4×(5)	f + u	流暢, (独創)
3. 用途テスト(用途)	8分	6×(6)	f + u	柔軟, (独創)
4. 問題点発見(問題)	5分	5×(5)	f	柔軟
5. 共通点発見(共通)	5分	5×(6)	f	柔軟
計	32分		創造点	

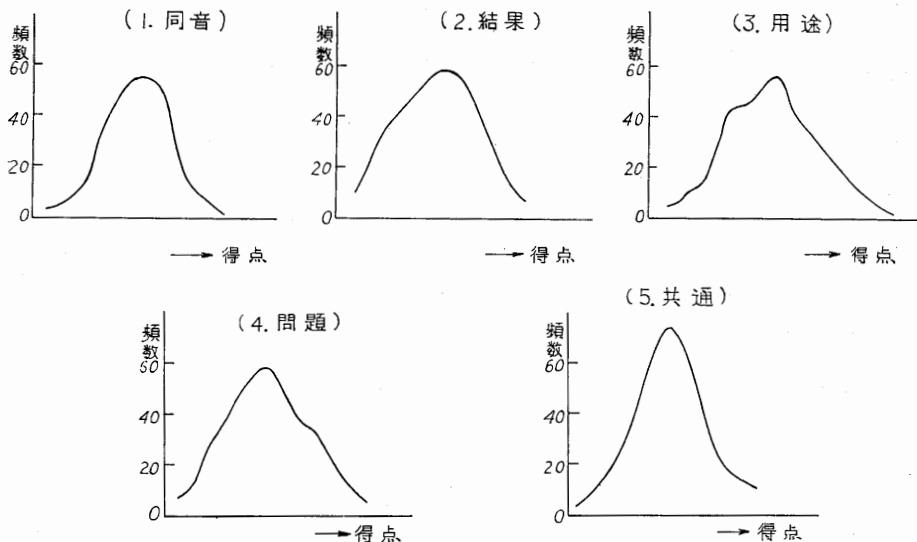
(注) f…正答カテゴリー数の偏差値

f + u…非凡反応カテゴリー数+正答カテゴリー数の偏差値

4. 標準化の基礎資料

資料は基準通りに採点, 集計された。また, 各下位検査ごとに偏差値を求めた。各下位検査の粗点分布は Fig. 1 にみられるように概ね単峰の正規に近いものであった。各下位検査別の平均, 標準偏差が Table 2 であり, 内部相関は Table 3 である⁹⁾。

Fig. 1 下位検査別粗点分布



9) 250名の学部別内訳は 文 23, 教 8, 法 33, 経 11, 理 16, 工 81, 医 31, 薬 18, 農 29 で, 現役112名1浪100名である。この250人と後述の290人全部を使用した場合の粗点平均のあいだには少し相違があるが, 有意差はない。

性差 本尺度における性差について検討をおこなった。

従来の資料から創造性テストの性差に関する確実な情報が全く得られないので、一般的傾向を確認する目的で次のような比較をおこなった。

(I) 標準化群と同じ学校の女子学生。これは最初からその目的で男子の標準化と同時に資料を採った。男一女成績を比較する。

Fig. 2 創造点得点分布

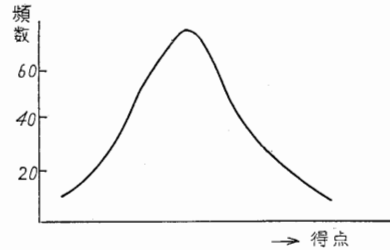


Table 2 粗点の下位検査別平均, S D

	粗点平均	S D
1. 同 音	43.43	7.3
2. 結 果	14.57	4.9
3. 用 途	14.81	5.8
4. 問 題	10.93	3.8
5. 共 通	12.56	3.0

N=250

Table 3 下位検査間内部相関表

	同 音	結 果	用 途	問 題	共 通
1. 同 音		.27	.22	.40	.16
2. 結 果	.27		.46	.57	.28
3. 用 途	.22	.46		.48	.38
4. 問 題	.40	.57	.48		.51
5. 共 通	.16	.28	.38	.51	
全 体	.54	.77	.75	.84	.76

(注) 男子106名の任意抽出

(II) 某公立大学の男女学生。(I)とはほぼ同じ時期に実施し、同様の整理、比較をおこなう。この集団の全数における男女比はほぼ1:1であり、知能程度・学力程度は(I)に較べ劣ると考えられる。

Table 4, Table 5 は男女の比較結果を示したものである。両表を対照すれば、創造点平均で約20点も京大が優り、下位検査成績にもレベルの相違がみられる。

男女比較の結果、創造点平均が京大で約8点、某大学で4点、いずれも女子が高いが、この差は統計的に有意なものでなく、男女差は全体としてみられない。

下位検査別に検討すると、(2. 結果)(4. 問題)の下位検査には両表ともに差がみられない。(1. 同音)は女子が優れていることは2集団の比較に共通

Table 4 京大(標準化資料)における男女差比較 $P < .01^{**}$
 $P < .05^{*}$

N	粗 点 平 均 (非凡点含まず)					偏差値計
	同 音	結 果	用 途	問 題	共 通	創造点
男子 290	42.25	14.23	15.40	10.81	12.01	249.7
女子 40	46.95	15.45	18.60	10.83	9.78	258.0
t	** 3.626	—	** 3.146	—	** 3.935	—

Table 5 某大学における男女差比較 (粗点及創造点)

N	同 音	結 果	用 途	問 題	共 通	創造点
男子 62	36.91	12.55	17.00	9.94	8.34	228.5
女子 165	42.07	12.90	14.98	9.66	8.81	232.5
t	** 4.458	—	* 2.249	—	—	—

してよみとれる傾向である。

流暢性能力に女子が高い傾向は Guilford¹⁰⁾ も予測しているが、この資料においても同様の傾向があらわれているといえる。

女子の得点の高さが平凡反応の差にあるのか、男子用の尺度であるために非凡反応に差があるのか検討するために別々に比較してみると、非凡反応には有意差はなく、したがって女子の反応自体が多いためと判断できる。

(3. 用途) (5. 共通) の両検査では2集団に一致した傾向がなく、何とも云えない。

参考迄に女子における内部相関を示しておく。

Table 6 女子における内部相関 (N=40)

	同 音	結 果	用 途	問 題	共 通
1. 同 音		.23	.45	.16	.31
2. 結 果	.23		.38	.24	.52
3. 用 途	.45	.38		.31	.31
4. 問 題	.16	.24	.31		.59
5. 共 通	.31	.52	.31	.59	
全 体	.71	.65	.89	.59	.71

学部別 学部ごとの人数に相違があることを考慮にいれる必要があるけれども、学部別の平均点を示すと Table 7 に示された結果を得た。経済学部12名の成績が他学部より若干良いことがわかる。

創造点平均の結果を、合格者の学力最高点、最低点の学部順位と比較してみよう。各学部の最高点合格者の学力成績と創造点平均の順位相関は $\rho=0.184$ 、最低点とは $\rho=0.034$ 、したがって全体として平均の創造点との対応は低く、学部別の学力検査結果との関係はみられないようである。

文理系別 言語性検査では文科系の人にくらべて理科系の人々には不利でないか。普通高校生にくらべ工業高校生の能力は低目に出るのはなかろうか。このような疑問に答えるために290名を文理系別にわけて創造点の平均を求めた。しかし、全体・下位検査ともに全く差がなかった。言語性検査だからといって、理科系の人に不利になるようなことはみられないことが示された。

Table 7 学部別成績と抽出比

学 部	抽出数	在籍数	抽出比	平均創造点
文 学	23	205	.11	250
教 育	8	50	.16	220
法 律	43	273	.16	240
経 済	12	220	.05	267
理 学	18	195	.09	246
工 学	94	795	.12	253
医 学	33	91	.36	254
薬 学	18	80	.23	253
農 学	31	210	.15	253
他 学 年	(10)	—	—	—
計	(280)	(2,119)	(.13)	(250)

高校卒業年度別 現役、浪人による創造点の相違は小さい。現役と浪人のあいだにあきらか
な差をみるができなかった。僅かに2浪の平均点が低い程度である。その他の分類に含まれ

10) Guilford, J. P. Merrifield, P. R. Cox, A. B. Creative thinking in children at the Junior High School levels. Rep. Southern California No. 26, 1961.

るのは検定合格者などである。

地域環境別 被験者に対し現住所（家庭所在地）と主として生育した地域について、市部、農村地域、漁村地域といった分類で答えさせた。しかし、現在のよう
に町村合併が盛んであり、都市と農村の境界がはっきりしなくなった時代では、区別がつきにくく判断が
任意的で一貫性を欠くと考えられる。郊外や混合地区と記入されたもの、あいまいなものを除き分類すると、主として成長迄の時期を六大都市の旧市域ですご
した64名の創造点平均は250点、農村地域57名248点、漁村地域3名236点となっている。其他の場合は分類
について考える必要が有るため省略する。

親の職業別 被験者は学生であって、彼等から職業別の分類をおこなうことができない、そこで質問紙に
記入された家庭の職業別に各個人を分類しようとした。分類にあたっては主観判断をまじえず記入された
通りとし、両親の職業が異なる場合には父親の職業で分類した。同種の職業が3名以上になるものの平均創造
点が Table 10 に記されている。その他の者は2名以下、無職、無記入、不明などである。職業分類には
社会学、産業心理学等の立場から考えられたものがあるがここでは参考迄に列举するに止める。

その他 きわめて受験成績が優秀と云われている高校（洛星、灘、北野、大手前）4校の出身者26名の創造
点平均は257.6点と全平均に較べやや高いが、特にすぐれているとは云えないように思えた。

5. 信 頼 性

各下位検査の信頼性は折半法を用いて算出した。分散の影響を異にする2種の公式にあてはめて算出した
結果が Table 11 である。各検査の信頼性は.5～.8で、

Table 8 文理系別検討結果

	文	理	t 検定
1. 同 音	43.47	41.73	—
2. 結 果	13.92	14.37	—
3. 用 途	14.21	16.04	—
4. 問 題	10.17	11.08	—
5. 共 通	11.47	12.24	—
6. 創 造 点	245.1	251.6	—
人 数	86	204	

Table 9 現役、浪人別創造点

種 別	N	平均創造点
現 役	128	251.55
1 浪	117	251.11
2 浪	20	246.15
3 浪	7	252.57
検定その他	8	219.60
他 学 年	10	—

Tabl 10 記入された家庭の職業別
創造点平均

職 業	N	創造点平均
大学教授, 教官	5	279.60
団 体 職 員	4	260.25
業 局 関 係	8	258.63
教 員	42	257.24
工 業 関 係	13	256.23
会社員, 銀行員	6	252.47
商 業 関 係	26	250.12
医 師	14	245.29
手工業, 家内工業	13	244.54
国家・地方公務員	36	244.50
自 由 業	6	243.50
農 林 業	16	240.44
サーヴィス業	3	236.33
食 品 関 係	5	234.80
その他(件数2以下)および不明	93	—

- 11) 増田幸一, 広井甫 創造テスト作成の試み 日教心第7回大会発表抄録 1960 P. 144—145. ここには
MHCT 創造性テストの中学生における折半法信頼度. 62高校生における再検査(5ヵ月)信頼度因子Iにつ
いて .38 IIについて .56 を報告している。我国における創造性テストの信頼性として本研究以前に公表
された数少ないものといえるものである。

比較的問題項目数が少ない点を考慮にいれば、ほぼ満足できる数値と云えよう。

全検査の信頼度推定にあたり、上記の資料から下位検査毎に奇数番項目、偶数番項目の得点をそれぞれ偏差値化し、その総計の値で以て信頼度を求めたところ .752を得た。再検査は日時の都合で折半法に用いた資料の全員に実施できなかったが50名で1年後に .805を得た。

この値いが創造性検査の信頼性として十分な高さであるかどうかは我国におけるこの種の研究が皆無に近い現状では断定できかねるが、被験者層は異なるが高校生・中学生における増田・広井の研究¹¹⁾に較べても遜色ないと云えそうに思える。各下位検査の係数値については Guilford¹²⁾の資料にみられる類似した検査の値に近いことが知られた。

女子に関する信頼性の値いは、男子にくらべ〔3.用途, 4.問題, 5.共通〕の各下位検査で幾分低くなっている。

6. 妥 当 性

妥当性を測る方法は極めて多く、一定していないが、この研究では3つの場合について妥当性の検討を紹介してゆきたい。

(a) 測定内容は知能と区別されているか

以前の研究は、創造的能力の発揮に一定水準以上の知能が必要であることが知られている。知能がある水準をこえた被験者では創造性テスト成績との間に低い相関がみられるとされる。また、紙とエンピツによる検査方法では一見して知能検査と区別できるとは限らない。そのために、本検査は知能とは別の、一定能力を測定しているかどうかの検討から始める必要がある。

因子的に知能、創造性を区別する目的で創造性各下位検査に知能検査（京大 NX15）を加えて相関を求め、完全重心法により寄与分散の大きい2因子を求めた。この2因子を直交回転した結

Table 11 折半法による信頼性の推定
N=106（標準化資料より）

テ ス ト	Mosier 公式 r	Rulon 公式 r
1. 同 音※	.724	.793
2. 結 果	.580	.736
3. 用 途	.754	.757
4. 問 題	.729	.715
5. 共 通	.505	.496
全 体	(折半法) N=106 r = .752 (再テスト法) N=50 r = .805	

※印のものは N=185 の値。

Table 12 女子における折半法信頼性係数
(N=40)

	Mosier 公式 r	Rulon 公式 r
1. 同 音	.780	.783
2. 結 果	.753	.745
3. 用 途	.122	.267
4. 問 題	.247	.246
5. 共 通	.277	.369
全 検 査	.700	

Table 13 回転後因子行列

	I	II	h ²
1. 同 音	.31	.60	.4561
2. 結 果	.66	.14	.4552
3. 用 途	.63	.03	.3978
4. 問 題	.79	.09	.6322
5. 共 通	.59	.02	.3485
6. 知 能	.10	.58	.3464

12) Guilford, J. P. Wilson, R. C. Christensen. P. R. A factor analytic study of creative thinking (II. Administration of tests and analysis of results) Rep. Southern California No. 8 1952 など。

果が Table 13 である。因子 I は創造性検査 5 種に共通して高く負荷し、第 II の因子は知能と (1. 同音) に高く負荷する知能因子であることがわかった。(1. 同音) の下位検査は流暢性の有効な測定器具である一方、平凡反応の量が知能と関係あることから、この検査がしばしば知能の検査としても用いられる¹³⁾ことが知られている。したがって、知能の因子にも高い負荷を示すのは予じめ予測されたところである。

このように因子を創造性と知能とに分離することができたが、これは下位検査を単位とした結果である。そこでもっとこまかく各項目別の妥当性をみる必要があるだろう。

通常の項目別分析にはある能力の上位、下位群による G-P 分析が用いられる。本検査の場合には創造性の上位、下位群間の弁別性検討では不十分でこれらの項目が知能と別の能力を測っていることの証明が必要であろう。

Table 14 項目別妥当性吟味のための 3 群

このため、Table 14 のような条件で知能、創造性の水準を異にした 3 群を設定し、相対的に項目分析をおこなった。もし、各項目が創造力をよく弁別し、知

群	知能偏差値	知能平均	創造点	創造点平均	N
知能 (I ⁺)	68以上	70.1	225点以下	215.8	20
創造 (C ⁺)	63以下	61.3	275点以上	280.7	20
劣等 (C ⁻ I ⁻)	63以下	59.2	225点以下	210.3	20

能の影響を排除できるならば、各項目での各群の平均得点は

創造群 (C⁺) > 知能群 (I⁺) ≥ 劣等群 (C⁻I⁻)

となるであろう。創造力レベルに関しては知能群と劣等群はほぼ等しく、知能テストは創造群、劣等群は同程度の成績しかあげ得ないと考えてよい。

結果を項目別に集計し、2 群ずつ対にして比較した。資料の煩雑を避けるため項目単位で、平均に統計的な差のみられるもの、平均に相違はあるが統計的には有意でないもの、の項目数を数え表示するに止める。

最初に知能一定な創造群 (C⁺) と劣等群 (C⁻) の両群を比較すると、Table 15 のように (1. 同音) を除き C⁺ > C⁻ の差は有意である。(1. 同音) の項目の弁別性は小さいものが多い。これは他の 4 種に較べて困難度の低い、項目数の多い検査であるためやや速度検査化し、一般の知能検査にみられるような、G P 分析においてテストの前半項目で弁別が小さくなる傾向があることによると思われる。しかし、平均点の同点もしくは逆転 (C⁺ ≧ C⁻) している項目はなく、全体として弁別性を示すものと考えてよい。他の 4 検査については問題がない。

Table 15 創造性にすぐれた群 (C⁺) と劣った群 (C⁻) の項目別平均点傾向比較

	項目数	C ⁺ > C ⁻ 有意項目数	C ⁺ > C ⁻ 非有意項目数
1. 同音	25	5	20
2. 結果	4	4	0
3. 用途	6	5	1
4. 問題	5	4	1
5. 共通	5	5	0

13) たとえばビネー・テスト、京大 SX 知能テスト、など。

創造群 (C^+) と知能群 (C^-) の比較においては、やはり全体として高い弁別の状態を示していると考えられる。(1. 同音)においては部分的に($C^+ < I^+$)となった項目も存在するが、この逆転は統計的に有意でない。(3. 用途)における($C^+ > I^+$)の有意でない3項目の差は、いずれも殆んど有意に近いものである。

Table 16 創造群(C^+)と知能群(I^+)の項目別平均傾向比較

	項目数	$C^+ > I^+$ 有意項目数	$C^+ > I^+$ 非有意項目数	$C^+ < I^+$ 非有意項目
1. 同音	25	1	16	8
2. 結果	4	3	1	0
3. 用途	6	3	3	0
4. 問題	5	3	1	1
5. 共通	5	5	0	0

(5. 共通)検査はビネーテスト等の知能検査に採用されているので、項目の選びかたが適切でない場合には知能の影響をうけることが心配されたが、この結果の範囲からは問題なく思われる。

知能群 (I^+) と劣等群 ($C-I^-$) はともに創造性の成績は劣るので、両群間の差はすなわち知能の量の影響である可能性が強く、理論的にはこの差が出現しないことが望ましい。

Table 17 創造性は同程度に低い知能群(I^+)と劣等群 (C^-) の項目別平均傾向比較

	項目数	$I^+ > C^-$ 有意項目数	$I^+ > C^-$ 非有意項目数
1. 同音	25	2	23
2. 結果	4	0	4
3. 用途	6	1	5
4. 問題	5	1	4
5. 共通	5	0	5

Table 17 の結果は ($I^+ > C-I^-$) の有意な項目は非常に少なく、(1. 同音)ですら2項目にすぎないことを示している。しかし、($I^+ > C-I^-$), ($I^- < C-I^-$)の項目数では若干前者の例が多い。

以上、総括して、この検査は各項目ともに大体において知能を分離して創造性を測定するものであると云ってよいように思える。

(b) 論理的基準との妥当性

知能との区別に関する分析が終ったところで、次に創造性と云われているものをこのテストはたしかに測っているか、という問題にふれなければならない。この証明法としてよくおこなわれるのは創造性が高い人は理論的に云ってどういった行動をするだろうか、といったことから、評定者がある一群の人々の創造的能力を論理的な基準にてらして評定し、その結果とテスト成績の関連をみることである。この方法はきわめて直接的、総合的であり、テストや質問紙での結果にはあわせえない多くの利点をもつが、評定者の主観が混入する度合が強く、かつ被評定者について、評定者が真にどれだけよく知りえているか疑問である。それに、われわれの標準化に使用した人達のような大学生では、彼等個人々々を的確に評定できる立場の人を求め得ない。

それならば、被検者自身に創造性検査とは独立に自己の能力その他を自己評定させるのも適当であろう。論理的に考えて創造的行動と考えるものについての自己判断と、実際の創造テスト成績との関連を調べた。

具体的には、妥当性検討のための項目を他目的の多くの項目に random に混入した。被験者も、できるだけ創造性に関する知識を持たないものが適当と考えて、標準化群を使用せず、この

標準化群の次年度の同じ大学の学生（このグループは、この種の研究に一切被験者として使用していない）を使用した。

被験者に昭和39年5月に創造性検査を実施した。教示においては単に知的能力のテストの一種と教示した。成績その他の情報は一切知らせず、約6ヵ月後、このテストの印象がほぼ白紙になったと思われる頃に質問紙を実施した¹⁴⁾。この中には、創造性とかなり関係があると思われる行動傾向を8項目まじえてこれを key item と呼ぶことにした。

創造性検査の結果から上位(G)群 20名、下位(P)群 20名を選んだ¹⁵⁾。

質問項目より key item を拾いあげ、その項目になされた各自の評定をあらかじめ定められた創造性の方向に則って5～1点を与えた¹⁶⁾。

GP 群別の評定平均点が Table 18 に示されているが、どの項目もわれわれの仮定した得点は

Table 18 創造性に関する Key Item と評定平均

	G 群	P 群	仮定との一致
3. 勉強の方法は (5) 自分の気ままにするのが好き (1) 他人に指導してもらった通りにするのがよい	4.65	4.15	+
17. あなたは、何か特許のとれるような作品または工夫をしるといわれると (5) 自信がある (1) 自信がない	3.35	2.50	+
22. シャレなどを云うのが (5) 好き (1) 嫌い	3.85	3.55	+
24. 新聞などにクイズが出ていたら (5) 必らず解いてみる (1) そのままにしておいて見向きもしない	2.65	1.95	+
25. あなたは自分について、大学生としては (5) 創造力のあるほうだと思う (1) 創造力がない	3.15	2.75	+
31. あなたは自分が知能テストと創造性テストをうけたとすると大学生の中ではどちらの方が成績がよいと思うか (5) 創造性テストの方がよい (1) 知能テストの方がよい	2.90	2.55	+
39. 何か問題をとくとき、次のどちらの型の問題が好きか (5) 正答のはっきりしていない—たとえば家をつくるときのアイデアなど (1) 正答のはっきりしている—たとえば代数をとくこと	2.40	2.15	+
75. 自分は全体として応用力のある人間だと思いますか (5) 思う (1) 思わない	3.50	2.90	+

(注) 上例では(5)が仮定した創造性の方向。但し、実際の質問紙では 方向は random において実施したが、結果をわかりやすくするために作表の際方向を統一した Table 19, 20, も同じ。

- 14) 質問紙は学習習慣傾向検査（このテストについては 日本心理学会第29回大会 1965 で報告した）と名付けるもので、自己の行動傾向や各種の好み-主として研究や勉強に関係あるもの-を 両極的な直線尺度上に5件法で評定するものである。全80問。たとえば物理と化学ではどちらを好むか。

A
B
C
D
E
 物理 (どちらでもない) 化学

- 15) 対象とされた人数100名で平均、分散はさきの標準化群（1年上級）と有意な差はみられない。

- 16) 最も創造傾向の強いもの5点。中間3点。として5～1の5段階配点。

G>Pとなった。検定は分布や人数の点で問題があるので実施しなかったが、とくに(3, 17, 24, 75)の各番号では0.5点以上の大きい開きがある。そして、テストの創造点の高い人々は、自分を平素から応用力があり、何とか自分でいろんなことを工夫できると考えていることが資料にはっきりあらわれて非常に興味ぶかい。また、相対的にはあるがこの人々は自分に創造力があると考えており、知能テストよりも創造性テストの方で自分は高い点をとるだろうと考える。これらの結果は、テストの妥当性を間接的に支持すると思われる。

質問紙の他の項目に関する詳しい結果の紹介は他の機会に譲ることにしたいが、そのうち一、二のものをとりあげてみよう。創造的な人間はものごとに打ちこむ、熱中する傾向があるとよく云われる¹⁷⁾が、精神集中度に関連する項目をぬきだして key item と同様の整理をおこなったのが Table 19 である。6項目中4項目には熱中、集中の傾向がみられたが、(33, 34)の両項目ではやや逆の結果がみられた。(33, 60)の結果は互いに少し矛盾するように思えるが、両項目の質問の意味は異なるのかもしれない。

Table 19 仕事や態度の集中傾向に関する item と評定平均

	G 群	P 群	仮定との一致
2. ある勉強をしている最中に (5) 他のことは考えない (2) 他のことをいろいろと考える	2.85	2.20	+
4. 長い小説などをよむときは (5) 一気に読んでしまう (1) 何度にもわけて読む	3.40	3.25	(+)
33. 何か、ことを始めると (5) 時間がきても熱中してやめられない (1) いつでもやめられる	3.40	3.70	(-)
34. 何か、考えごとに集中するまでに (5) すぐに集中できる (1) 時間がかかる	2.70	2.75	(-)
60. 仕事をしながら (5) 時間のことは忘れてしまう (1) 時計をよくみる	3.00	2.60	(+)
68. ある問題について考えているときは (5) そのテーマについて集中的に日夜考えぬく (1) とくに集中的に考えることはしない	3.75	3.10	+

計画性 (Planning) に関する項目では両群の示す傾向のあいだに大きなちがいがあることが示された。Table 20 の結果は、G群は計画やスケジュールをたてるのを好み、整理には図表を多用し、結論を直観的に求める傾向が強い。理解のしかたとか方法とかいったものは人間には好みがあるから一概に有益な方向を予かじめ定めておくことはできないが、両群のあいだにこのように大きい開きがみられることに注目したい。

(C) 実際の妥当性の例

(C 1) 大学入学後の進級状況と

17) たとえば Van Zelst. R. H. & Keer. W. A. Some correlates of scientific and technical personnel, J. appl. Psychol. 38, 1954, 145-147.

Table 20 計画性 (Planning) に関する item と判定平均

	G 群	P 群
13. ある複雑な文章を理解するには (5) 図表にしてまとめることが多い (1) 文にしてまとめることが多い	3.85	3.15
18. ものごとを考えると (5) 直観的に考えることが多い (1) 因果関係的につきつめていく	3.50	2.50
38. 旅行や試合などのスケジュールを (5) 決めるは好き (1) 嫌い	3.45	2.75
69. 何かプランを考えると (5) エンピツとペンで書きとめながら考える (1) 頭の中で全く考える	3.05	2.80

標準化群の大学入試成績と創造性検査成績、大学への高点合格者とその知的能力、などについては別の文献¹⁸⁾で詳しくのべた。要約すれば、合格者のみにしてみれば学力検査との関係は創造性よりも知能の方が僅かに高いこと、但し高点合格者は普通合格者に較べて知能、創造性ともに高いことなどが明らかにされた。そこで、次に大学入学後の実績との関係を調べる必要があるが標準化群はまだ卒業していないので成績としては教養部から専門課程への進級成績が唯一のものである。しかし、この成績を各人について調べるのは困難であり、たとえ調べたとしても各人の課目選択の組合せが多様であって、また成績評価にも教師差が考えられる。したがって詳細にしてもこれらの理由からの確かな判断はなされにくい。そこで、単純に専門課程への進級可否判定を基準として用いることにしたい。

進級の可否は¹⁹⁾進級、仮進学、留年に3大別される。不合格者(仮進学プラス留年)の中には疾病、海外渡航、政治活動などあきらかに学力以外の理由からのものが含まれるので、判明する限りこれらの資料は除いた。

標準化群はこの及落判定時期の1年半前に創造性検査をうけている。われわれ被験者中から留年11名、仮進学28名を拾いあげ不合格群とした。

対照群として進学者中から、さきの不合格者群と学部別に人数が等しくなるようにして²⁰⁾39名を random に抽出した。

18) 住田幸次郎 大学入学試験における学力検査成績についての分析 京大教育学部紀要 Vol.10 77-93 1964

19) 専門課程進級の条件は語学、自然科学、社会科学、人文科学、体育などにわけて定められた最低単位数をすべて満たしていることである。この単位数は、学部により少しことなるところもあるが、殆んど同じと考えて処理しても誤りはないと思われる。

(i) 進級 単位のすべて揃っているもの。

(ii) 仮進学 原則として、1課目分の単位が不足する程度するとき、1年以内にその課目に合格することを条件に一応専門の聴講を許されるもの。ここでは事実上不合格者として扱う。

(iii) 留年 単位が不足で(ii)の仮進学の条件にも満たないため留級にされるもの。昭和37年入学京大生概数約2,000名中仮進学は約300、留年100、ほぼ全体の20%にわたる。

20) たとえば、不合格者中に法学部の者が3名含まれていたとすると、資料中の法学部の進学者全員から3名を任意にえらんだ。

21) 知能偏差値は合格者平均 66.3 点不合格者平均 64.9 点で $P > .05$ (t 検定) である。したがって差は有意でない。

両群は知能偏差値²¹⁾に統計的有意差なく、入試成績もほぼ等しかった。

Table 21 は、両群の創造点平均を比較したものである。合格者群は不合格者群にくらべ有意(1%)に得点が高い傾向がみられた。この傾向は留年、仮進学と個別に比較しても同様である。創造点平均は統計的に有意差はないが留年群と仮進学群とではむしろ留年群が高い。

この結果について一つの解釈としては、留年者の中には学力以外の原因で単位が多量に不足している者があるのに比べ、仮進学群は合否すれすれのところで能力的に不合格になったのかもしれない。この点についてはより詳細な分析が必要であるが差は統計的に有意でない故に、ここではこのくらいにしておきたい。

合格者群は全体として創造点平均が高いことはわかったが、つぎに創造点のレベルとの関係を調べよう。創造点平均250の±30点(280~220)より両極端にはいる人数を調べると、上位は合格者群12名(30.8%)に対し不合格群は5名(12.8%)と、両者の含まれる人数はかなり異っている。逆に、低い方の極端には2

名、8名と不合格者が多い。

Table 22 はこれらの結果を示したものである。

以上の諸事実は、不合格者に

Table 22 進級可否と創造点の程度

群	N	創造点の範囲	280点以上の人数(%)	220点以下の人数(%)
合格者	39	331—215	12 (30.8%)	2 (5.1%)
不合格者	39	344—189	5 (12.8%)	8 (20.5%)

比べ進学者の創造点が高いことを示すと共に、創造力が大学における学習適性に関与する可能性のあることを示していると云えよう。但し、注意すべき事は、創造点の高い不合格者の存在すること、両群の創造点の分布範囲は殆んど同じであること等から、創造点の程度と進学可能性の関係は一義的に決定できず、他の要因の介在を考える必要があろう。しかし、創造性得点の高さ、低さは進級を規定する一要因としては考えるべきであろう。

(C2) 現場への適用例

実際に現場で働いている人々への適用例を示してみよう。会社幹部として活躍している人々は、この種の創造性検査にどのような結果を示すか考察するのも妥当性検証の手段として有効であろう。某中堅製薬会社の協力を得て、同社の課長、次長クラス19名に本検査を実施した。彼等の年齢は35~40才、全員が旧制の大学、専門学校の卒業生で文科系7名、理科系12名である。実施は昭和38年8月、実施者は人事部長である。その結果、19名の創造点は315~204点に分布した。平均は253.8点と大学生の標準化群に較べても劣らない。ただ、Fig. 3のように下位検査別の成績にムラがあり、主として柔軟性を必要とする検査では低下がみられる。しかし、自分の経験の豊富さを活かせると思われる検査(たとえば4.問題)では、きわめて高い得点を示す。彼

Table 21 進級可否と創造点平均
**P<.01

群	N	創造点平均	
合格者	39	264.7	
不合格者	39	243.9	**
留年	11	(246.4)	**
仮進学	28	(243.0)	**

等の結果を調べてみると、大学生にくらべて反応内容に相当著しい相違があり、大学生の平凡反応と特殊反応の中間あたりの性質の反応がこれらの人々には多いように思われる。平凡反応は彼等が意識的に避けるのではなく、彼等と大学生の迎った経験の質的量的な差および年齢の昇により適切反応つまり平凡反応をとりだす機敏性の低下等知能テストの年齢による衰えと同種のものがあらわれているのではなかろうか。

彼等の仕事上の能力を評定することは、職務内容の相違や立場の相違から比較が困難であったが、19名全員をよく知っている関係者により総合的にみて極めて優秀と考えられる2名が選ばれ、彼等の創造点を調べたところ2名とも300点を越える良い成績であった²²⁾。

19名を出身校の面から大まかにすると Table 23 のようになり、従来の一般的評価にはほぼ一致する傾向を得た。

其他の結果については評定に個人的な問題やハロー効果が予想されるので省略する。

(参考 I) 高校生、中学生への適用

本尺度は大学生により標準が作成されている。

高校生については、大学生とほぼ同様に考えて適用することが可能と思われる。優秀高校の3年生に本検査を実施したところ、概要次のような結果を得た。

(i) 粗点平均は標準化大学生にくらべ等しいか、またはやや低い。したがって創造点も低目に出るが、個人的には大学生と同等または優秀な者もある。分散は大学生にくらべ大となるが、これは標準化の集団の知的能力がかなり高いレベルにあることから当然であろう。

(ii) 高校生100名から作成された標準による創造点と、大学生の創造点の相関を高校生43名について求めたところ、両者のスコアは.97と極めて高い一致をみた²³⁾。

(iii) 高校生についても、いくつかの実際の妥当性の例証を得た。

以上のように、高校生については大体そのまま適用できると思われるが、中学生については今後の研究を待ち、その結果により中学生向きのものを作成する必要があるだろう。

Fig. 3 会社幹部テスト成績の標準化群との対比

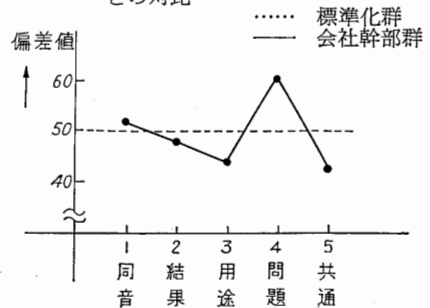


Table 23 会社幹部の出身校と創造点

	人 数	創造点
京 大	7	273.4
他 の 国 立 大	2	237.5
公 立 大	1	241.0
京 都 の 私 立 大	3	228.7
薬 専	3	237.3
旧制専門 (技術関係)	3	270.7
計	19	253.8

22) 二人の創造点は315点と306点であった。

23) 両スコアの得点絶対値の一致度が高いことを意味するものではなく、相対順位の一一致が高いことを示すものであることは云う迄もない。

(参考II) 本検査尺度についての発表資料

1. 梅本堯夫 住田幸次郎 創造性検査としての同音異義連想検査の研究 関西心理学会第71回大会(大阪市立大)発表資料 1962
2. 住田幸次郎 大学入試と知能・創造性テスト成績に関する一研究 関西心理学会第72回大会(奈良女子大)発表資料 1963
3. 住田幸次郎 テスト法による創造性の測定に関する基礎研究 心評, 7, 211—228 1963
4. 住田幸次郎 テスト法による大学生の Creativity 測定の試み 日本心理学会第27回大会(日大)発表記録 P.128 1963
5. 住田幸次郎 大学入学試験における学力検査成績についての分析 京大教育学部紀要, 10, 77—93 1964
6. 住田幸次郎 独創反応の出現と時間経過の関係についての研究 関西心理学会第75回大会(甲南大)発表資料 1964
7. 梅本堯夫, 城戸幡太郎, 苅阪良二, 住田幸次郎 生産性と創造性に関する研究 知能と創造性の関係についての分析 日本教育心理学会第6回総会(九大)発表抄録 136—137 1964
8. 住田幸次郎 創造性テスト成績と日常の学習習慣傾向の関連 日本心理学会第29回大会(青山学院大)発表抄録 P.174 1965