

認知型テスト日本版に関する一研究*

布留武郎・渡辺 良・佐賀啓男・生田孝至

はじめに

学習の個人差を規定する新しい要因として、ここ10余年来アメリカの心理学者によって認知型に関する研究が行われてきた。認知型 (cognitive style) というのは、個人が情報に接した際、まず情報のカテゴリゼーションを行ない (知覚レベル)、符号化 (encode) された情報は蓄積され (記憶レベル)、内部的な仕上げをうける (思考レベル)。この一連の情報処理の仕方は、接した情報の性質によって異なるけれども、類型的な個人差があると考えられている。ごくおおまかにいえば、情報を全体的に処理する様式と分析的に処理する様式と2つの型にわかれる。東京教育大学の辰野グループはこの研究に注目し、アメリカで開発された認知型テストのうち、Witkin の EFT、Kagan の CST および MFF の3者を取りあげ、その日本版の開発を試みている。

われわれはマス・メディア特にテレビとマンガへの接触が児童の認知型の形成に一役を買っているのではないかと考えた。児童が日常接するテレビ番組はその大部分が娯楽番組であり、マンガ本やマンガ雑誌と並んで、これら絵画メディアへの接触 (以下テレビとマンガを含めて絵画メディアと呼ぶ) は、全体的・文脈的認知傾向を促進する一方、分析的・熟慮的認知傾向を阻止する方向に機能するであろう。この仮定を支える1つの研究として、辻功の「へき地児童に与えるテレビ学校放送の効果」(昭38)があげられる。それは当時まだテレビのなかった群馬のある地域に、実験校と対照校を設け、実験校には「テレビの旅」と「テレビ理科教室」とを1年

* 本研究は放送文化基金の援助によるものである。

間継続利用させて、その効果を学力テスト等によって測定したものであるが、学力を統制するために実験期間の前後にB式知能検査を施行している。その結果、予想外の事実が発見された。それはテレビ番組の継続視聴によって知能得点が上がるということである。下位検査別に実験校と対照校を比較すると、精神速度あるいは記憶の因子負荷量の高い単純な作業において有意差があらわれたが、熟慮的分析的思考を要する複雑な作業には差がみとめられなかった。この事実は研究者が指摘しているように、認知テンポが標準より遅いへき地児童が、標準的ペースで進行するテレビ画面に適應することによって、精神速度や記銘がはやく正確になるためではないかと推察される。

絵画メディアにおいても、そのプログラミングの仕方によって、分析的認知を促進するような番組を作ることは可能である。しかし児童が日常接触するテレビ番組やマンガ本ではそのようなプログラミングは行われていない。学校放送の「テレビの旅」においても同様であり、「理科教室」も低学年番組は別として、高学年番組では20分という時間的制約もあってと思うが、分析的認知の発展を目的とする番組は作られていない。児童が頻繁に接触する絵画メディアはストーリーがあって、画面の構成や展開は文脈的関係的であり、展開のペースは省略技法によってかなりはやくなっている。画面のテンポがはやくなっているというのは筆者の印象にすぎないが、「のらくろ」時代のマンガと今日のマンガ、初期のテレビドラマと今日のそれとを比較するならば、客観的証拠が得られるであろう。筆者が1974年のプリジュネースに参加した際の印象によると、ヨーロッパ諸国の青少年向けテレビ番組に比べて、日本やアメリカの作品はテンポが顕著にはやい。これは筆者と共に参加したNHKのディレクターと共通の印象であった。

このように絵画メディアへの累積的接触は児童の認知傾向に一定の方向を与えつつあると考えられるが、今日のようにテレビが普及し、マンガ本やマンガ雑誌がいつでも児童の手の届くところにある社会では、メディア

接触の効果を接触群と非接触群とに分けて比較することは不可能だから、結局、重接触者と軽接触者との比較に終わってしまう。したがって仮説のような効果があっても、フィールド調査ではその効果があらわれにくいと考えられる。

そこでわれわれは本調査を実施する前段階として、約200名の小学5年生男女に対し、認知型テストとマス・メディア接触の調査を行ない、両者の関係を見ることにした。認知型テストは辰野グループが開発しつつある3種類の日本版を下敷として、そのテスト項目の検討から作業を開始した。本報告では認知型テスト作成の作業とその結果に限って述べる。

認知型テスト作成の準備作業

EFT

Embedded Figure Testの略字で、Witkinにより開発された。一般用のテストは簡単な幾何図形を複雑な図形の中から探し出すもので、作業を困難にするために複雑な図形は標準図形とちがった形に彩色されている。色が干渉して標準図形を認知することがよりむずかしくなるわけだ。児童用として幾何図形の代わりに、具象的事物の略画を使用したテストも作られているが、辰野グループの日本版は、児童用として、原版の複雑な幾何図形を若干簡易化して、彩色を省略している。それだけ原版に比べて困難度が軽減されている。テスト項目は25題あって、集団テストが可能であり、正答の多いほど、場独立的 (field independent) な認知傾向がつよく、その逆は場依存的 (field dependent) な認知傾向がつよいと判定される。

場独立的とは、状況を分析的にとらえ、対象を背景から分離して認知する傾向を指し、場依存的とは、状況を全体的枠組でとらえる傾向を指す。場依存型は発達的に場独立型に先行する。人間の心理機構における個人差は分化の程度により生ずるもので、その知覚領域でのあらわれが、このような認知型であるという考え方に基づいている。

EFT は、RFT (Rod and Frame Test) のような測定のために心理学実験室を必要とするテストから発展した簡便に実施できるテストで、身体的定位感覚は含まれていないが、RFTとの相関が高く、また安定性も高い。ただし、RFTとちがって練習効果がはっきり認められる。Witkinらによって、知能、記憶、学業成績、性格等との関係が研究されている。十分に一致した結果は認められないが、興味ある発見は、非人間的課題(単語暗記や変形図形命名など)の記銘は場独立型の方がまさり、社会関係を含む課題(人の顔の再認)では場依存型の方がまさるということである。また、パースナリティとの関連において両型を区分する傾向がみとめられたことは大いに意義のある発見であろう。それは永年心理学者の課題であった知覚領域とパースナリティ領域とを結合する接点として、重要な意義を荷う可能性があるからである。

わが国における追試研究はここでは省略して、われわれがEFTの集団用日本版を使用するにあたって、模様替えを試みた点を次に列記する。

(1) 練習問題2題は辰野版では線図型であるのを、面図型に改めた。テスト問題は面図型であるから、作業の転移がより容易になるであろう。

(2) 一部の複雑図形に線を補足して認知を容易にした。辰野版には、線を補足しないと標準図形を形成しない複雑図形が二、三あった。

(3) 問題に難易のあることが一見してわかるので、やさしいものからむずかしいものへと、らせん型の配列法をとることにした。問題の困難度は大学院生11名とA教会の日曜学校児童18名とに調査を行ない、だいたいの見当をつけた。

制限時間は辰野版では6分となっている。しかし、院生の場合でも25題に反応するのに5分半を要したこと、また、日曜学校児童の場合、6分間では反応数が最大23(6年女)、最小5(2年男)で、その平均は25題中14題(正答数はさらに少ない)だったので、制限時間を7分にすることに決めた。

C S T

Conceptual Style Test の略。Kagan の開発したテストで、枠の中に人物、動物、道具類などの略画が三つ描いてあるセットを呈示して、類似しているもの、あるいは一緒にしたらよいと思うものを二つ選ばせる。分類の仕方には、分析的・関係的・推論的範ちゅう化の3通りがある。分析的とは、刺激を視覚的に分析し、要素の類似性に基づいて分類することであり、関係的とは、刺激間の機能的関係に基づいて分類することである。推論的とは、刺激間の共通な性質を推論し、それに基づいて分類することである。テストは20ないし30セットによって構成され、分析的反応を多くするものを分析型、関係的あるいは推論的反応を多くするものを非分析型と呼んでいる。

テストの安定性は少なくとも10才以後は高く、発達的には非分析型が先行する。知能その他の知的側面や学業成績等との関連が研究されているが、わが国における研究を含めて、断定できるほどの一貫した結果はでていない。しかし、C S T と知能の因子である動作性尺度と正の相関があるとか、分析型と非分析型とは概念学習のストラテジィに差があるとか、両型の相違は能力にあるのではなく、刺激のどこに注目しやすいか、換言すれば、刺激に対する構えにあるとかなど、認知型という概念構成をうらづける結果が出ていることは見逃せない。

C S T の集団用日本版には辰野版のほかに、多田や新城が特定の実験計画を遂行するために試作したものがある。われわれは、それらを下敷にしつつ、次の点を考慮にいれて改訂版を作成した。

(1) カテゴリゼーションが部分の類似性によったのか、機能的関係によったのか、推論にもとづいたのか、あいまいな刺激の対は避けること。参考にした上記のセットには、この点のあいまいな刺激の対がかなり見うけられた。ただし採用したセットの中にも、この点を追求すると論理的に割り切れないものが含まれている。例えば、紳士と腕時計と物指の描いてあるセットでは、腕時計と物指を選んだ場合、両者の目盛の類似性による

分析的認知反応と判定される。しかしひょっとすると、計器という上位概念を媒介とする推論的認知反応であったかもしれないのである。集団テストではそれを確かめることができないので、準備段階における個別調査の経験に基づいて常識的に判断した。

(2) 分析的・非分析的の反応がいずれか一方に極端に偏よることを予め防止するため、絵柄や絵の配置を考えること。参考版の中には、絵のある部分が特に目立って、反応の大多数が“分析的”にかたよることが一見してわかるような絵や、部分が目立たないで、文脈的關係が言語連合によって密着している絵があった。後者の場合は、たとえ分析的反応をする割合が僅少であっても、困難度の高い項目として使用できるが、前者の場合は、“分析的”というより、絵の形態的類似に注目した反応というべきかもしれない。一般に属性による分類は機能のそれに先行するプリミティブな心理機構とされているが、対象の形や大きさや色などに共通性を認知するメカニズムと、対象の細部に注目して類似性を発見するメカニズムとは、同じく属性分類の中に含まれるとしても、発達的に区分されるべきだと思う。CSTにおける認知型の分類に、属性・機能・クラスといったカテゴリーを導入すると混乱をもたらす。Kaganの“分析的”とは絵の細部に注目するという意味で使用されていることはあきらかである。

(3) 推論による分類に問題のあること。推論的範ちゅう化とは、刺激間の共通な眼に見えない性質を抽象し、それに基づいて分類することであるならば、この定義に見合う絵を描く必要がある。例えば、同形同大同定位の2枚の木の葉の絵が並んでいて、その一方の一部が欠損しているセットにおいて、この2枚の絵を選んだ被験者は、はたして木の葉という上位概念を媒介にして、両者を結合したのであろうか？ それとも単に全体的形態の類似性を認知した故にであらうか？ 1枚の葉は一部が欠損しているけれども、他の葉と同じものであることは眼に見えているのであって、目に見えない共通の性質を抽象するといった推論のプロセスをふまなくては分類できないのであろうか？

同様な疑いを抱かせる絵のセットが多く、上位概念の媒介を必要とするものは、むしろ例外的に少なかった。それ相応の理由があるのだろうが、われわれは定義に忠実な操作をするために、推論的範ちゅう化の対刺激については、低次の上位概念の媒介を必要とするように改めた。例えば、2枚の葉は樹木の種類をちがえるとか、2人の少年は一見して同一人物でないことがわかるようにするとかした*。

上記の原則にしたがって、26種類のセットを作成し、集団調査が可能なように1頁に3セットを配置した。26セットの配置順はランダムに行なった。そしてB日曜学校の児童・生徒19名に対して予備調査を行ない、分析的反応対非分析的反応の比率を求め、それを参考に試行錯誤をくりかえしつつ、一つの試作版を作成した。既成のセットに比べて、ほとんどどのセットも絵の細部の認知に、ある程度の熟慮 (reflection) を要するようになっている。

M F F

Matching Familiar Figures の略。Kagan はC S Tの分析反応には、視覚的分析傾向と熟慮性との2因子が含まれていることを発見し、問題解決学習には後者の因子が非常に重要であると考えて、熟慮性—衝動性 (Reflection—Impulsivity) を測定するテストの一つとして、M F Fを開発した。それは誰にもなじみの深い対象の略画を使用し、標準刺激と同じものを6つの類似刺激中から選ばせる作業で、正答に達するまで選択作業を行なわせる。問題は12題あって、測度として、第1反応 (正誤を問わない) までの平均時間と、総誤答数を取り、反応時間が中央値以上かつ誤答数が中央値以下のものを熟慮型と呼び、その逆を衝動型と呼んでいる。

両型は問題解決のストラテジィの相異を示すもので、アイカメラを使用

* あとで気がついたことだが、“推論”という用語にこだわりすぎたようである。若干修飾された刺激の同一性を認知するという意味で、同形・同大・同定位の対の結合もまた推論的範ちゅう化といえるかもしれない。

してMFFテスト実施中の眼の動きを観察したところ、熟慮型は選択図形間の差異に注目し、衝動型はその類似性に注目する傾向があったということである。知能その他知的側面との関係をしらべた研究では、熟慮型は文字や語の誤認が少なく、読書能力にまさること、帰納的推理を要する課題で誤まりが少ないこと、一方、衝動型は不確定場面で質問回数が多く、拡散的思考にまさることなどを明らかにした研究もある。また、衝動型には一般的活動性の高いものが多く、熟慮型には運動の言語的統制にまさるものが多いことを示す研究もある。

辰野らの日本版では、絵の素材は略画辞典からとった21組のセットから構成されている。この課題は作りやすいので、困難度をさまざまに変えることができる。改訂版を作るにあたって注意した点は次のとおりである。

選択肢中、標準刺激とちがうことが一見してわかるものはせいぜい2つ以内にとどめる。一見してちがいがわかる選択肢は選択肢としての価値はないが、一方、回答者のテスト不安を減少させる効用があると考えられる。

この原則に従って辰野版から10セットを選定し、若干修飾を施した。これに Kagan の標準テストを構成する12セットを加え、計22セットをA・Bの2系列に折半して、反応時間と選択肢の誤答分布をみるための準備的調査を行なった。A系列は Kagan の11セットを含むテストで、被験者は小学3年生から5年生までの男女を含む26名、B系列は辰野版を中心とするテストで、被験者は小学2年生から中学1年生までの男女を含む41名（うち34名はA系列と同じく3年生から5年生）であった。いずれも個別テストで、その結果、(1) 第1反応時間の平均値が逸脱的に高いセット、(2) 正答に達するまでの反応時間の平均値が逸脱的に高いセット、(3) 選択肢の誤答分布が極端に偏るセット、(4) 第1反応の誤答が極端に少ないセットを除くことにして、A系列から8セット、B系列から4セット、計12セットを選ぶことにした。(1)と(2)の規準は、困難度が高すぎて被験者に不安を与え無気力感に陥ることを防止するためであり、(4)は問題が易し

すぎて弁別性を欠くおそれがあるためである。(3)は選択肢の重みに偏りがあり、一部の選択肢に誤答が集中すると、他の選択肢はなくてもよいことになるからである。

予備調査の結果

町田のある公立小学校5年生5学級、205名の男女児童に対し、1974年11月のある数日間にわたって、テレビ・マンガ本・マンガ以外の本・ラジオ・レコード等、マス・メディア接触量を中核とする質問紙法による調査と、前記3種類の認知型テストを行なった。

実施の手順は次のとおり。第1日：EFT, CST, マスコミ調査の順に、午後の1時限を割いてもらい、3学級122名に集団で実施。この調査と平行して、残りの2学級80名(3名欠席)に、MFFを実施。このテストは7名のインストラクターにより、1時限1学級の割合で個別に行なった。第2日：MFFを施行した2学級中の1学級に対し、午前中の1時限を割いてもらい、EFT, CST, マスコミ調査の順に実施。第3日：残りの1学級に対し同上の調査を同様の手順で実施。インストラクターはいずれも当大学院視聴覚教育専攻の学生であった。

したがって、EFT, CST, マスコミ調査は5学級205名のデータが得られ、MFFはそのうちの2学級80名のデータが得られたことになる。その結果の概要は、認知型テスト自体の検討、マスメディア接触量とその内容、両者の関係の3部分にわけて報告する予定だが、ここでは認知型テスト自体について述べる。

EFT

制限時間7分間。できる問題からさきに手をつけるよう教示した。

得点分布は平均値13.00(25点満点)、標準偏差4.66、範囲は1~25で、全問に正答したものが1名あった。このテストを調査第1日目に行なった3学級122名の男女に限定して、各項目の誤答数、無反応数をしらべたと

表 1. EFTの項目別反応分析
(Table 1. Response analysis of EFT)

項目 (item)	誤答数/反応数 (number of errors/number of responses)	誤答率 (ratio)	無反応数 (no-response)
(1)	2/120	2%	2
(2)	17/121	14%	1
(3)	8/115	7%	7
(4)	1/120	1%	2
(5)	19/122	16%	0
(6)	31/117	26%	5
(7)	30/114	26%	8
(8)	35/117	30%	5
(9)	33/116	28%	6
(10)	12/99	12%	23
(11)	51/106	48%	16
(12)	10/107	9%	15
(13)	27/92	29%	30
(14)	25/101	25%	21
(15)	16/92	17%	30
(16)	11/55	20%	67
(17)	17/60	28%	62
(18)	21/52	40%	70
(19)	11/41	27%	81
(20)	10/40	25%	82
(21)	10/26	38%	96
(22)	2/15	13%	107
(23)	9/16	56%	106
(24)	6/11	55%	111
(25)	3/8	38%	114

N=122

ころ、表1の結果を得た。

表1の示すように、誤答率は不規則ではあるが困難度が低高・低高と、らせん状に分布している。また、無反応数は10題目あたりから顕著に増し、16題目からさらに増大し、以後漸増する。時間制限テストとして典型的な分布といえよう。このようにテスト得点の散布度及び項目別反応分布

から考えて、EFTは修正を加えずに本調査で使用することにした。

CST

このテストにおける反応の特徴は、EFTのような能力テストとちがって、分析的あるいは機能的、分析的あるいは推論的と、少なくとも2通りの反応が可能なことである。したがって、教示の仕方が反応に影響を与える。Kaganの原版では、3つの絵の中から“似ている”絵、あるいは“いっしょにしたらよい”絵を2枚えらぶように指示を与えるだけである。したがって被験者が“どちらにしようか”と迷った場合は、被験者の判断に任せられている。辰野版では、“初めに思ったものを選んでください”という指示が与えられるので、被験者が2通りあるいはそれ以上の回答の仕方があることに気づいた場合でも、回答は初発反応に限定される。いずれの教示がこのテストの趣旨に合っているかについてわれわれの間で討論した末、原版の教示にしたがうことにした。初発反応に限定する方が、テスト実施中に起こるかもしれない被験者の質問を予め防止することができる点で都合がよい。しかしそうすると、分析型の特徴である熟慮性のインディカントを消してしまうおそれがある。絵刺激自体が統合的認知方向へ圧力がかかっているようなセットでは、分析型の被験者でも、初めに気づくのは統合的關係であるかもしれないからだ。事実、機能的に關係する絵のセットには、「マッチとタバコ」「ペンとインク」「キュウスと湯呑み茶碗」「フォークとナイフ」「蝶と花」といった強度の言語連合ができあがっていると思われる対が多いのである。このような刺激の対に対して、非分析型は衝動的に言語連合による回答をするけれども、分析型は別の分類方法もあることに気づいて、どちらにしようかと考えるだろう。その故に、Kaganが発見したように、分析型の反応時間は非分析型のそれより長いのである*。

* 初発反応に限定すると、刺激の部分に注目するか、全体に注目するかという認知のスタイルがより純粹にあらわれるかもしれない。教示の与え方については研究の余地がある。

このようにCSTの反応は分析性と熟慮性の2因子を含み、また反応は絵刺激の内容、部分の描き方、3枚の絵の位置など、刺激自体の要因によって規定されることが多いので、標準化することは非常にむずかしいと思われる。開発者のKagan自身も、のちにはCSTから離れてMFFに傾斜するのであるが、CSTは集団検査が可能であり、認知能力でなく認知の構えを測定する尺度である点に魅力がある。

以上の見地に立って作成したCSTは26題からなり、うち13題は分析対クラス、13題は分析対機能の2通りの分類ができるように仕組まれた。項目の順序はランダムである。制限時間は設けず、26題終了までにだいたい5分間あれば十分であった。

5学級205名についての分析得点の平均値8.54、標準偏差4.80、範囲は1~20で、分析反応の困難度はかなり高い。

初日にこのテストを施行した3学級122名について、上位得点群40名、下位得点群40名をえらび項目分析を行なったところ、表2の結果を得た。項目の弁別性はGuilfordの ϕ 値により、Jurgensenの計算図表を使用した。

分析反応100分比の範囲は3%~70%、平均35%で、項目別にみても非分析反応に偏っている。これは課題の性質上、分析反応の困難度が高くなるように設計したからである。分析反応が10%以下の(2)、(4)、(8)、(20)はいずれも非分析反応が機能的言語連想として定着していると考えられた項目である。 ϕ 値が5%水準に達しない(7)と(17)は非分析反応がクラスによる範ちゅう化を要求される項目である。

Kuder-Richardson 20式による信頼性係数 $r_{tt} = .89$ で、項目の内的一貫性は高い。

本調査用のテストとして、弁別性を欠く(7)、(17)を除き、新しく「トンボ、縞模様のストッキングをはいた少年、蝶」の1セットを加えることにした。分析反応が10%以下の上記4つの項目は、弁別性があるので採用することにした。

表 2. CST 項目分析
(Table 2. Item analysis of CST)

項目 (item)	分析反応比率 (ratio of analytic rersponse)	弁別性 (ϕ)
(1)	20%	.33
(2)	9%	.32
(3)	61%	.23
(4)	8%	.31
(5)	19%	.28
(6)	15%	.28
(7)	33%	.14 (n. s.)
(8)	3%	.18
(9)	38%	.47
(10)	56%	.16
(11)	37%	.43
(12)	25%	.28
(13)	69%	.32
(14)	66%	.55
(15)	24%	.56
(16)	32%	.19
(17)	34%	.14 (n. s.)
(18)	24%	.29
(19)	20%	.22
(20)	9%	.20
(21)	70%	.29
(22)	59%	.37
(23)	41%	.60
(24)	49%	.59
(25)	35%	.46
(26)	49%	.52

N=122, 但し ϕ 値は上位群40名, 下位群40名をデータとしている。 ϕ 値の
 p. 05=.177
 p. 01=.233
 p. 001=.299

M F F

練習問題2題は Kagan の標準テストと同じものを使用，また絵刺激の大きさ，選択肢セットの大きさ（たて21cm，よこ24cm）も Kagan の標準テストと等しくした。1枚のセットの中に6個の略画が2段に並んでいる。標準刺激の位置は選択肢セットの左側中央におく方法と，上部中央におく方法とあるが，Kagan 同様，後者を採用した。テスト項目12題はランダムに配列，インストラクターと被験者は対面に座し，選択肢を人さし指で押えるように，また正答に達するまで反応を継続することを教示した。反応時間（第1反応）はストップウォッチで計測した。所要時間は1人あたり5分間あればよいことが再確認された。

このテストは教室使用の都合上，当学級の教室で，7名のインストラクターが手わけして平行的に行なったので，被験者の注意が相互に干渉をうけたこと，インストラクターの院生が面接に習熟していなかったために，面接者個人差があらわれたこと，また，各項目ごとに正答に達するまで反応を求めたので，テストの終り頃には飽和状態に達する児童が見うけられたこと，この3点のハンディキャップを考慮に留めておく必要がある。

2学級80名のデータによる全12項目に対する第1反応平均時間は12.3秒（SD=6.4），範囲は4.1秒～37.4秒であった。項目分析を行なった結果は次のとおりである。

(1) 反応時間

第1反応を行なうまでの時間の平均値を項目別に秒単位で示したのが表3である。

項目(6)と(8)は，反応時間が60秒をこえる児童が若干名いるが，反応時間の項目平均値，その散布度を指標とする限り，テスト項目を修正する必要はないと思われる。

(2) 誤答分析

項目別に選択肢の誤答分布をしらべたのが表4である。

さきに述べたように，選択肢作成の原則として，一見して標準刺激とち

表 3. MFF 反応時間
(Table 3 MFF : Response time)

項目 (item)	平均値 (M)	S D	最大値 (MAX)	最小値 (MIN)
(1) 電気スタンド K (lamp)	9.49 sec.	4.94	23.50	1.00
(2) 蝶 (butterfly)	11.97	7.85	56.00	2.90
(3) 家屋 K (house)	11.00	6.37	33.00	2.20
(4) 蛙 (frog)	12.07	8.61	51.60	3.80
(5) 電話器 K (phone)	13.72	7.02	42.30	4.00
(6) キリン (giraffe)	13.51	10.65	80.60	2.80
(7) ドレス K (dress)	13.70	8.97	48.00	2.60
(8) 船 K (boat)	14.50	10.84	80.00	2.90
(9) テレビ受像機 (TV set)	13.56	9.78	54.00	3.20
(10) 椅子上の熊 K (bear)	12.45	7.96	45.70	2.20
(11) はさみ K (scissors)	9.59	5.05	27.90	2.40
(12) 猫 K (cat)	11.85	10.74	56.90	3.50

N=80, K=Kagan から採用した項目

表 4. MFF 選択肢誤答分布
(Table 4. MFF : Distribution of Errors)

項目 (item)	選択肢 (alternative)						計 (Total)
	①	②	③	④	⑤	⑥	
(1)	3	21	19	4	0	17	64
(2)	10	0	10	22	7	11	60
(3)	0	9	10	13	16	7	55
(4)	5	33	23	0	13	11	85
(5)	9	18	0	4	11	1	43
(6)	25	8	25	6	27	0	91
(7)	13	28	23	11	0	2	77
(8)	15	0	10	8	7	6	46
(9)	10	10	3	0	23	19	65
(10)	0	26	0	17	9	2	54
(11)	1	2	12	6	4	0	25
(12)	28	7	0	12	11	8	66
計 (Total)	119	162	135	103	128	84	731

N=80, 0は正答選択肢にあたることを示す
(0 shows a correct alternative)

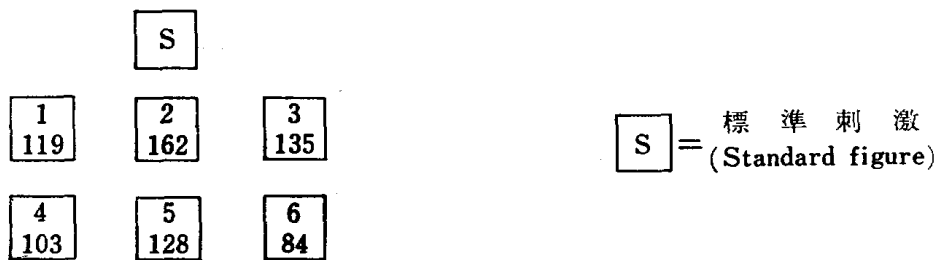
がうことがわかるような選択肢は2つ以内にとどめるという規準を設けた。この規準の誤答率を仮りに5%ないしそれ以下におくと、 $N=80$ だから誤答数4乃至それ以下の選択肢が3つ以上ある場合は、この規準からはずれることになる。項目(1)はそれに該当するが、残りの選択肢も誤答が少なく、バランスを著しく失うほどでもないので、抵抗の小さい易しい課題として、残しておいてもよい項目である。他の項目はすべて、上記の規準をみたしているので、どの項目をとっても実質的に最小4つの選択肢は確保できることになる。

また、項目別誤答数の計(表右端欄)をみると、順番があとになるにつれて、誤答数が増大するというような傾向は認められない。したがってこの限りにおいて、被験者の飽和を示す証拠はない。

次に、標準刺激の位置と選択肢配置の関係から生じる反応バイアスについてみると、図1の示すように、中央に位置する選択肢②と⑤は、それぞ

図 1. 選択肢の位置と誤答度数分布

(Figure 1. Arrangement of alternatives and error's frequency)



れ左右の選択肢より誤答数が多く、また上段の選択肢は下段の対応する選択肢に比べて誤答数が多い。このように位置関係による反応バイアスが存在することは否めない。標準刺激を左側中央においた場合といずれが反応バイアスが小さいかは残された一つの課題である。しかし反応バイアスの存在が誤答数の個人差測度に重大な影響を及ぼすとは考えられないので、

さしせまった問題ではない。考えられるのは、衝動型は熟慮型に比べて刺激の布置による反応バイアスを受けやすいということだが、若しそうだとすると、反応バイアスの因子は衝動型の誤答数を多くするように作用するので、むしろ熟慮型との選別性を高めるものといえよう。

MFFテストは、正答に達するまで選択反応をくりかえすことを要求している。したがって、弁別のできない被験者はフラストレーションをおこすおそれがある。事実、そうした児童がいたことが報告されている。またテスト所要時間はできるだけ短いことが望ましい。こうした見地に立てば、誤答数も第一反応だけに限った方がよいことになるが、一方、そうすると誤答数の散布度が小さくなるおそれがある。そこで反応を何段階までとれば、第6反応までとった場合と散布度がほぼ等しくなるかをみるために、反応段階ごとの誤答累積度数の平均値と散布度を示したのが表6である。

表 6. 反応段階別誤答累積度数の平均値と散布度

反応段階	平均値	SD	最大値
第1反応	5.95	2.22	11
第2反応まで	8.26	3.35	16
第3反応まで	8.98	3.87	18
第4反応まで	9.08	3.97	19
第5反応まで	9.12	4.06	20
第6反応*まで	9.13	4.10	21

N=80

*第6反応では誤答の最大値が増加しなくなるのが普通だが、1人の被験者は第6反応において、前に選んだ誤りの選択肢を再び選んでいるため20から21に増加している。

表6によると、当然のことながら、第5反応までとれば、全反応をとった場合と誤答数の平均値と散布度は殆ど変わらないことがわかる。第3反応で打切っても大差なく、テストの実用性から考えると、誤答の測度に関する限り、正当に達するまで反応を求めず、第3反応あたりで止めた方がよいように思われる。

(3) 反応時間、誤答数の相関

各被験者ごとに項目別の反応時間、その計、及び反応段階別誤答数、その計について相関行列を算出したところ、次のことがわかった。

反応時間（第1反応）を項目別にみると、どの項目をとっても、12項目を合せた測度と高い相関がある（.59～.90）。また項目間の相関係数はすべて5%以下の有意水準に達している。

したがって、反応時間に関していえば、12項目全部について測定しなくても、テストの信頼性は保留できると思われる。

誤答数を反応段階別にみると、どの段階においても累積度数の測度と全段階を合せた測度間には次に示すように高い相関がある。

第1反応	.8919
第2反応まで	.9616
第3反応まで	.9921
第4反応まで	.9982
第5反応まで	.9996
第6反応まで	.9999

すなわち、第3反応で打切っても、全反応の誤答数分散の98%をカバーできるわけである。

反応時間と誤答数の相関係数は、Kaganにしたがって第1反応時間対全誤答数でとると、-.54である。両変数の関係を項目別、反応段階別にとっても、両変数を対に組合せたすべての相関係数は1%以下の有意水準を示し、-.54と大差ない。

以上の結果から、反応段階のみでなく、項目数を削減しても、テストの信頼性、項目の妥当性にほとんど影響を及ぼさないことがわかったので、実用的見地から、項目を10個に減らし、かつ誤答の測度を第3反応までの範囲でとることにした。厳密に言えば、どの項目を除くべきか、またそうした場合に第3反応までに限っても誤答の測度が変わりがないかをしらべるための分析が必要であろう。しかし実用的見地にたてば、いままでの分析データによって、12項目からランダムに2項目を抜いて、選択反応を第3

段階までに限定しても、テストの信頼性や項目の妥当性に影響することはほとんどないと考えられる。

念のため、2つの項目を除くための客観的データとして、誤答数及び反応時間の2側面から次の分析を行ってみた。まず第3反応までの総誤答数を基準変数として、項目別誤答数を説明変数とする重回帰方程式を求めたところ、項目(5)と(11)の β ウエイトが比較的小さいことがわかった。

項目	β ウエイト	項目	β ウエイト
(1)	.2176	(7)	.2139
(2)	.2566	(8)	.2122
(3)	.2331	(9)	.2126
(4)	.2446	(10)	.2190
(5)	.1999	(11)	.1329
(6)	.2745	(12)	.2639

次に全項目の反応時間平均値を基準変数とし、項目別の反応時間を説明変数とする重回帰方程式を求めたところ、項目(5)、(11)の β ウエイトは比較的小さいことがわかった。

項目	β ウエイト	項目	β ウエイト
(1)	.0640	(7)	.1161
(2)	.1015	(8)	.1405
(3)	.0824	(9)	.1266
(4)	.1114	(10)	.1030
(5)	.0909	(11)	.0653
(6)	.1378	(12)	.1389

この2つのデータから、項目(5)と(11)を除くことに決定した。さらに厳密には、残りの10項目について、反応を第3段階で打切った場合と正答に達するまで反応を要求した場合の誤答分布の比較をすべきであろうが、実用的にはそこまで行なう必要はないであろう。

3種類のテスト間の相関

EFT, CST, MFF-RT, MFF-ER間の相関をこれら3種類の

表 7a 認知型テスト間の相関 (全体)

Table 7a Correlations between Cognitive Style Tests (Total)

	C S T	M F F-R T	M F F-E R
E F T	-.0521	-.2023	-.0523
C S T		.1130	-.0722
M F F-R T			-.5434*
M F F-E R			

N=80 P.o5=.2209

表 7b 認知型テスト間の相関 (男)

Table 7b Correlations between Cognitive Style Tests (Boys)

	C S T	M F F-R T	M F F-E R
E F T	.1461	-.2461	.1905
C S T		.1636	-.2604
M F F-R T			-.7004*
M F F-E R			

N=46 P.o5=.2914

表 7c 認知型テスト間の相関 (女)

Table 7c Correlations between Cognitive Style Tests (Girls)

	C S T	M F F-R T	M F F-E R
E F T	-.3637*	-.1503	-.2858
C S T		-.0202	.1696
M F F-R T			-.3756*
M F F-E R			

N=34 P.o5=.3396

テストを併用した80名のデータによって求めたのが表7a, b, cである。男女を合せた表7aでは、前述したMFF反応時間(RT)と同誤答数(ER)間に有意な負の相関がみられる以外には、いずれの対にも5%水準の有意性もみとめられない。この結果は日本を含めた先行研究(CST対MFF, CST対EFT)と合致する。したがって、これらのテストは認知型に関する独立の因子を含むものといえそうである。

しかし男女別にしらべると、女子ではCSTとEFT間に5%水準の負の相関がみられる。両変数については、205名のデータがあるので、それによって両者の相関をみると、

全体： .0127

男子： .0573

女子： -.0849

で、女子においても有意でない。

したがって表7cにおける両変数の有意な相関は、少数サンプルによる偏りと解釈すべきであろう。ランダムイゼーションを欠いた小サンプルには往々にして、このようなシステムティック・バイアスが存在する。

あ と が き

EFT, CST, MFFが何を測定するものかについては、これらテスト間の相関のほか、知能、創造性、性格、学業成績等との相関を求めることによって、先行研究に新しい知見を加えることができるであろう。これらのデータについては、本調査の結果をまっせ漸次報告する予定である。

なお、筆者の眼を認知スタイルに向けてくれたのは新城君の修士論文であり、今回の研究にあたって同君所有のEFT, MFFの標準テストの現物を借用することができた。また辰野グループの認知型研究に関する展望はわれわれのよき水先案内となった。3種の認知型テストの日本試作版を入手できたのは同グループの小林幸子さんの好意による。これらの人々並びに調査に進んで協力してくれた院生の諸君に対し深い感謝を捧げる。

(布留武郎)

参 考 文 献

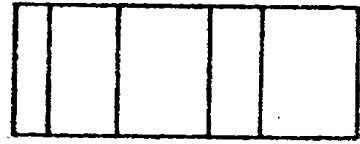
1. Goodenough, D. R. and Eager, C. J. (1963). A modification of the embedded-figure test for use with young children. *J. of Genetic Psychol.*, 103, 67-74.
2. Kagan, J., Rosman, B. L., Day, D. Albert, J. and Phillips, W. (1964). Information processing in the child. Significance of analytic and refl-

- ective attitudes. *Psychol. Monographs: General and Applied*, 578, 1-37.
3. Kagan, J. (1965). Reflection-impulsivity and reading ability in primary grade children. *Child Development*, 36, 609-628.
 4. Kagan, J. (1965). Reflection-impulsivity: The generality and dynamics of conceptual tempo. *J. of Ab. Psychol.*, 71, 17-24.
 5. Kagan, J., Pearson, L., and Welch, L. (1966). Conceptual impulsivity and inductive reasoning. *Child Development*, 37, 583-594.
 6. Kagan, J., Pearson, L., and Welch, L. (1966). The modifiability of an impulsive tempo. *J. of Ed. Psychol.*, 57, 359-365.
 7. 小林幸子・辰野千寿・沢田瑞也 (1972)。認知型テストに関する研究—C S T と E F T の検討を中心にして—, 教育心理学会第14回大会発表論文集。
 8. 小林幸子, 幼児における認知タイプの検討 (1972), 日本心理学会第36回発表論文集。
 9. Messick, S. and Damarin, F. (1964). Cognitive styles and memory for faces. *J. of Ab. & So. Psychol.*, 69, 3, 313-318.
 10. Satterly, D. and Brimer, M. A. (1971). "Cognitive styles and school learning". *The British Journal of Educational Psychology*. 294-303.
 11. 沢田瑞也・小林幸子・辰野千寿(1973), 認知型テストに関する研究(3), 第15回教育心理学会発表論文集。
 12. 新城岩夫 (1973), 認知スタイルの学習と教授法への影響に関する一考察, 国際基督教大学修士論文。
 13. 多田俊文 (1966), 子どもの認知タイプと視覚教材のみかたの関係について—C S Tによる実験報告—, NHK放送文化研究年報11, 47-69.
 14. 田上不二夫・根本橋夫・小林幸子 (1973), 衝動型・熟慮型テスト作成の試み—個別テスト・集団テストの比較—, 第15回教育心理学会発表論文集。
 15. 辰野千寿・福沢周亮・沢田瑞也・上岡国夫・小林幸子・高木和子・伊藤康子 (1972), 認知型に関する教育心理学的研究, 教育心理学年報第12集。63-107.
 16. 辻功 (1963), へき地児童に与えるテレビ学校放送の効果(3)—知能検査, 学力検査の結果を中心として—, NHK放送文化研究所年報 8。142-163.

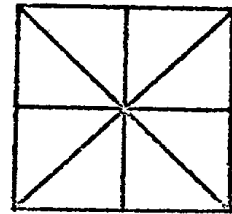
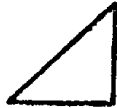
付 録

EFT

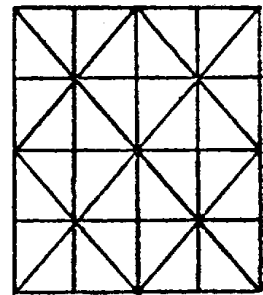
札んしゅう
1



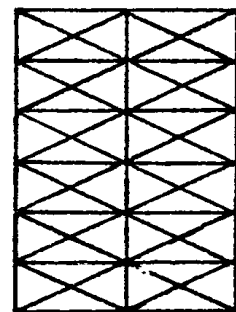
札んしゅう
2



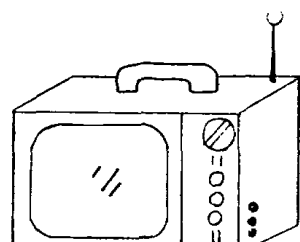
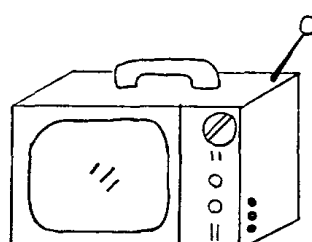
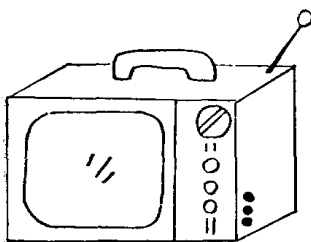
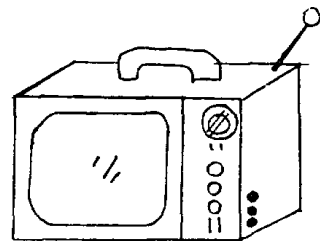
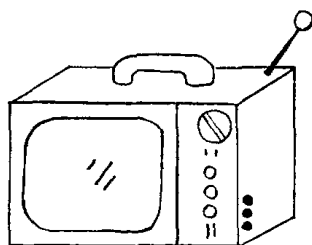
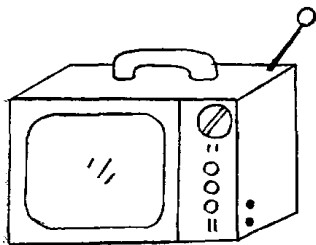
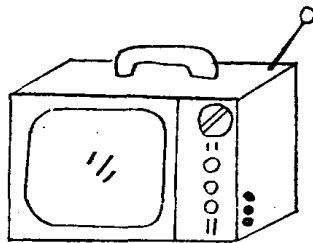
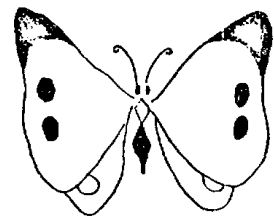
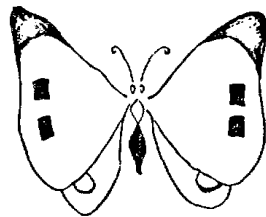
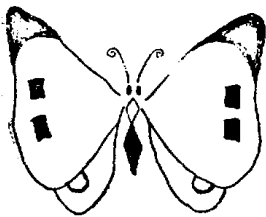
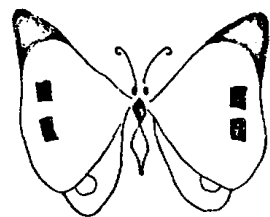
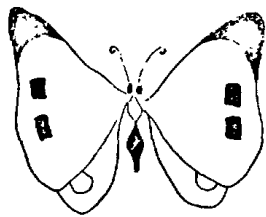
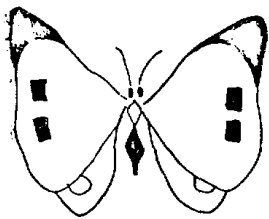
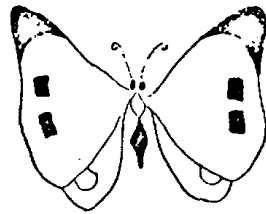
1



11



M F F



A Study on the Japanese Version of Cognitive Styles

Takeo Furu, Ryo Watanabe, Hiroo Saga
and Takashi Ikuta

We assumed that an accumulated exposure to TV and comics might have some effects on cognitive styles of children. Before confirming this assumption, we carried out a study to revise the Japanese versions of GEFT (Group Embedded Figure Test), CST (Conceptual Style Test) and MFF (Matching Familiar Figures) which were developed tentatively by the Tatsuno team, Tokyo Kyoiku University. The GEFT consists of twenty-five geometric figures which are employed from Witkin's Test and slightly simplified and coloremphasized for children version. The CST including twenty-six familiar figures and the MFF including twenty-one familiar figures are both similar to Kagan's Tests. Our principle to revise these Japanese versions and the findings by study are as follows.

GEFT

We rearranged the test items to make a spiral from due to their difficulties which were determined by a preliminary study, using eleven graduate students and eighteen children of a Sunday school. And we conducted the GEFT in the five classes of fifth graders at a public school located in the suburb of Tokyo metropolitan area, with the total number of two hundreds and five. The time limit of performance was seven minutes.

The mean score gained from this test was 13.00 (SD=4.66), and the range of scores was 1 to 25 (a girl gained the full mark). Table 1 shows the response pattern to each item. This analysis was made with the data gathered from three classes out of the

five, totaling one hundred and twenty two children. Response error ratios range from 1 % to 56 % and tend to form a spiral figure of difficulties, although they do not show regular distribution pattern. The number of non-responses tends to increase remarkably as items come later. The coefficient of reliability was $r_{tt}=0.88$ (Kuder-Richardson's formula 20).

CST

Responses to CST are supposed to be much influenced by the stimulus itself. We made a new Japanese version on the basis of following standards.

- (1) Three familiar figures in a set must be paired in the two ways, either 'analytical vs. functional categorization' or 'analytical vs. inferential categorization'.
- (2) A functional categorization would be likely confined to a simple relation between two stimuli, for example, 'a match box & cigarettes', and a pair implying some fantastic relation would be avoided.
- (3) Similar parts of some two figures which induce an analytical categorization should be drawn less remarkedly.
- (4) An inferential categorization should be made with a mediating concept.

The new Japanese version of CST including twenty-six items was conducted with total subjects. The instruction was the same as Kagan's. The mean analytical score was 8.54 (SD=4.80) and the range was 1 to 20. The result of item analysis (Table 2) with the same three classes also showed that responses were somewhat inclined to non-analytical direction (the mean of analytical responses by item was 35%, and the range was 3% to 70%). It is, however, better than the other case, because the test has been constructed to assess a tendency of analytical perception and a ceiling effect will not likely occur. The coefficient of reliability was $r_{tt}=.89$ (Kuder-Richardson's formula 20).

MFF

We selected arbitrarily ten items from the Japanese version, and added the Kagan's twelve items to those. Then we split those twenty-two items into two parts, A & B. After a preliminary study which was conducted with two groups of children assigned respectively to the form A or B, we have adopted twelve items out of the both forms on the basis of the following criteria: to eliminate any item showing (1) too high response time mean, (2) almost non-errors at the first response, and (3) error tendency to concentrate in a few alternatives.

We conducted the revised version of MFF with the two classes out of the five, totaling eighty boys and girls. The mean response time was 12.3 sec. ($SD=6.4$) and the range was 4.1 sec. to 37.4. Table 3 shows the means of the first response time by each item. Table 4 shows the frequency distributions of errors on alternatives by each item. Those which have too low error frequencies would not function as an alternative, while they would do as a facilitating response agent. From this point of view, almost all items seem to have substantially four alternatives at least. We can find that there is no tendency to indicate some sort of saturation, because the error by item does not increase as it become later (see the right column of Table 4). We can find, however, a response bias stemming from the arrangement of alternatives (Figure 1). The alternatives of no. 2 and no. 5 arranged at the middle have more errors than those at the left or the right, and alternatives on the first line have more errors than those on the second line. This sort of response bias, however, would function to differentiate responses on the two dimensions of reflection and impulsivity.

It is better to shorten performance time of the test without lowering reliability. We made a correlation matrix of response time by item, the total response time, accumulated errors at each response step, and the total error. Then we found that all correlations were statistically significant at the 5% level. The association between the

total response time and total error showed $r = -.54$. We found also a correlation between total error and accumulated error at the third response step was .99, in other words, 98% of the total error variance can be explained by the errors at the third response step. Besides, response time to any item had significant correlations with the errors at the third step. Therefore even if we stop responding at third step, the reliability of MFF would not be changed.

Relations among the Three Cognitive Style Tests.

Any correlation among the three cognitive style tests was not statistically significant (Table 7). This means they include independent factors.