

## 1. はじめに

平成2年度より3年計画で、「授業設計におけるコンピュータの高度利用」として、教育訓練や関連の授業での情報化社会に即応したコンピュータの活用方法の研究を行ってきた。‘授業’は計画、実践および評価の3つの側面で議論されることが多い。‘評価’は新たに授業設計を行うための客観的な資料を得るための重要な過程であり、本研究においては、コンピュータテストによる学習の評価法やアンケートデータの処理法を開発するなど、‘評価’を一貫して重視してきた。しかし、一般に教育訓練の評価は、カリキュラム、教材教具、技能など、対象が多岐にわたるものであり、本研究でとりあげた研究対象は、むしろピンポイント的な内容といった方がよい。本年度は、テストやアンケートの評価に関する前年度までの成果を踏まえてこれをさらに発展させたが、今回は最終年度であるので、より広く、教育評価の方法を資料としてまとめることを一つの主眼としている。

教育訓練現場で、評価のあり方を検討されるにあたり、その一助になれば幸いである。

## 2. 教育訓練の評価

### 2.1 評価の目的

教育評価は、教授学習活動の一環とみなされ、そのこと自体が目的というよりも、その後の授業計画や何らかの意思決定への反映を意図してなされる。その意図にあわせて、評価の手法や評価用具を選定し、評価モデルを構成する。したがって、評価に際しては、その目的を明確にする必要がある。授業・学習において、評価の目的は、およそ次のように分類される。

#### (1) 授業方略にかかわる意思決定

学習者の状態を教師の発問やテストなどを通して評価し、評価結果にもとづいて、その先に学習進行するかそれとももう一度反復するかを決めたりより易しい教材に変更したりするなどは、いずれも、評価をその直後の意思決定に利用していることになる。CAIにおいては、いかにその評価モデルを構築するかが課題であり、テスト構成法の研究につながることが多い。そして、テスト回答の情報をもとにして、最適な学習コースの分岐の仕方を決めるわけである。実際の授業では、テストや事中での質疑にもとづいて意思決定を行うことが多い。

## (2) 授業や教材の設計

講義の評価や授業中に使用した教育機器の評価は、その後の授業の進め方に関して、示唆を与える。また、教材を開発するときに、どのような学習内容において学習者がつまづきやすいかなどのデータが得られていると、やはり教材の設計に役立てることができる。

ここでは、テストまたはアンケートを主な評価用具とする。授業を評価するときはアンケートが多く用いられるが、評価対象が漠然としているため、発問項目は“授業”の考察に役立つような内容を精選する必要がある。

学力に関する教育情報は、これをコンピュータに蓄積しておくこと、学習状態を時間を追って分析することが容易になる。個別の学力を学習者全体と比較して分析することも可能になり、特に注意を払うべき学習者や再検討すべきテスト問題を抽出するのに役立つ。

## (3) 査定

学期末のテストは、所定の学力を取得したと見なされるかどうかの査定である。また、入学試験は、各教科の得点を求めたのち、一元化して合否を査定するのが直接の目的である。評価にある程度主観を伴うような論文テストでは、そのあいまいさの故に、二人以上の採点者が別々に評価することがある。しかし、論文テストと客観テストのいずれでも、一旦査定結果が得点に表されたのちは、いささかのあいまいさもなく、総合得点が計算され合否が決定される。

## (4) 学習活動の促進

特に、数学、物理などのように知識を基礎から積み上げ、系統的な学習を要するものは、途中で定着テストを幾度か行い、学習状態を確認することが多い。学習者の知識の定着を図り教授者の指導方法にフィードバックするために行うとの目的で実施されるが、それと共に、評価結果を学習者が認識することにより、その学習意欲を固めることができる。そもそも、頂点のありどころのまったくわからない山登りを強いられるほど、登山意欲を減退させるものはない。辿ってきた一連の下位の学習目的を図式で示し、全体の中のどの目標まで確実に歩んできたかを自己確認することが、さらに続く学習目標の達成のための原動力となり不要なフラストレーションを起こさせないものである。

## (5) 教育モデルの有効性の確認

ある教授技術を開発したり、新しいメディアを試作したときには、実際に授業の

中に取り入れ試行することによって、どのような効果が得られるか、また、どのような問題が残されているかを、実践に即して考察することが可能になる。ただ、教授・学習は、教師、学習者、学習内容など諸々の要因が関与する上、有効性なる概念自体あいまい模糊としているので、モデル授業に取り入れた教育メディアの有効性の有り無しが、一元的にまた2値的に帰結されることは殆どないといってよい。アンケートやテストといった評価用具を活用して評価データ得、さらには適宜教育研究者などの意見をも得て、様々な可能性を考察しつつ新たな授業設計の方針を定める。

実際の教育訓練現場において、評価は、多くの場合、上記の(1)~(5)のどれかに該当する。そして、より具体的な目的に応じて評価の枠組や基準を定めることになろう。

## 2.2 評価の手法

教育評価は、指導法、教材、学習者、教育システム、教育環境など、対象が多岐にわたる。実際の教育研究では、評価対象をどれかに絞るが、教授・学習を表2.1に示す評価要因でとらえ、当該評価対象をその中に位置づけると便利ことが多い。この表は、坂元昂氏の評価法を翻案・拡張したものである。そこに優先順位が記されているが、これは、たとえば新しい教育モデルを授業に導入するという場合に、2種類の教育モデルのうちどちらが適切かを比較して決めるために、次のように活用する。

いま、2つの教育モデルA, Bが、同表に占めされた評価要因で比較され、評価結果が数値化されているものとしよう。モデルAにおける学力の評価値を(学力)<sub>A</sub>などと表すことにする。もし、(学力)<sub>A</sub> > (学力)<sub>B</sub>かつ(情意)<sub>A</sub> > (情意)<sub>B</sub>となったら、これらが共に最優先項目であるので、他の項目の評価結果にかかわらず、総合評価値はモデルAのほうが高いと判断する。次に、(学力)<sub>A</sub> > (学力)<sub>B</sub>かつ(情意)<sub>A</sub> = (情意)<sub>B</sub>かつ(交流)<sub>A</sub> < (交流)<sub>B</sub>となったら、交流面より学力面のほうが優先順位が高いので、学力の評価比較が優先され、評価値はやはりAのほうが高いとの帰結を生む。さらに、(学力)<sub>A</sub> > (学力)<sub>B</sub>かつ(情意)<sub>A</sub> < (情意)<sub>B</sub>となったら、これらが同一の優先順位であることから、この限りでは決着がつかない。

評価要因に、時間面や経済面を加えた場合でも、評価の手法は上記とほぼ同様である。

ここで、2つの教育モデルA, Bを想定したとしても、同じ学習者に同じ学習内容を教えることはできないので、ふつうは多数の学習者を互いに等質な2グループにわけ、別々の教育モデルを割り当てるという方法を採用することが多い。

表2.1での5つの評価要因や評価用具を以下に簡単に説明しておく。

1. 「学力」は、学習者に与えた課題に対する応答から測定することが多い。また、アンケートを用いて評価を行うこともある。
2. 「情意」は、学習者に、授業に導入した教育モデルに対するイメージとして問うたり、意見として問う。
3. 「交流」は、よく小学校や中学校を対象としてとりあげるテーマである。
4. 「時間」は、一定の学習内容を修得するのに要する時間を意味するが、CAIではその測定がしやすい。
5. 「コスト」は、たとえば機械系で実機を教材として用いるなど大きなコストがかかる場合は、重要な要因となる。

実際の授業では、学力面は、個別に質疑を行いその回答内容から理解度を推測することが多いが、すべての学習者から同一条件で回答を得ることは困難である。情意面もほぼ同様であり、したがって客観的なデータを得るには、個別の学習者に対して同一の方法でデータ採取する必要がある。その点において、ペーパーテスト、コンピュータテスト、質問紙法によるアンケートは、より客観的な手法といえるわけである。

表2.1 教育訓練における評価要因

評価要因（優先順位）	評価用具、計測方法
学力面(1)	テスト（アンケート）
情意面(1)	アンケート
交流面(3) ・教師－学習者間 ・学習者相互	アンケート
時間面(4)	時計
コスト面(5)	金額

## 2.3 評価用具～テストとアンケート～

テストとアンケートのいずれも、評価の目的に応じて、評価内容を適正に選定し、問題の構成を図る必要がある。ここでは、これまでよく用いられているテストとアンケートの分類例を、問題例とともに以下に記載する（橋本重治 昭51）。

### 2.3.1 テスト

#### (1) 客観テスト

採点者によって評価が変わらないもので、割合断片的な知識をテストするのに適している。

#### (a) 再認形式

答を選択する方式であり、テストのコンピュータ移植が容易である。

#### 1) 真偽法

真か偽かを2値的に回答させるもの。

(例題) Prologは手続き型言語である。 1. 真 2. 偽

#### 2) 多肢選択法

設問に対し、正答を含むいくつかの選択肢を提示して回答させるもの

(例題) Lispは、次のどの処理を目的として開発されましたか。  
選択肢 1. 技術計算 2. 記号処理 3. 事務処理 4. オンライン処理

3) 2つの選択肢グループを見比べて、ある‘関係’を形成するペア構成を回答させるもの。

(例題) 左のグループの技術に携わった人を右のグループから選びなさい。


〈技術〉	〈人〉
1. コンピュータ	1. エジソン
2. 飛行機	2. フォンノイマン
3. 畜音機	3. ライト兄弟

(b)再生形式

単語や複合語程度の答を自由に記述させるもので、簡単な答ならコンピュータ入力が可能である。

1) 単純再生法

名称などの単語や簡単な複合語を回答させるもの。

(例題) 電子回路で、は何を示すか。

2) 完成法

文章の一部を空白にしておき、文脈などから推理してその部分を回答させるもの。

(例題) 次の文章で、箱の中に適切な用語を単語をいれて文章を完成させなさい。  
コンピュータのハードウェアは、大きく3つに分けて、中央演算処理装置、  
装置および装置から成る。また、ソフトウェアは、その中枢としてがあり、各種の処理プログラムはすべてその下で稼働する。

1) 訂正法

誤った部分を訂正させるもの。

(例題) 次の文章で、誤字に下線を引き、訂正しなさい。

「経済大国としての日本は、これまで開発登上国に対してさまざまな援助を行ってきたが…」

2) 配列法

事柄の順序を回答させるもの。

(例題) 次の3つの事柄について、年代の古い順に1, 2, …と書きなさい。  
( ) ソビエト連邦が崩壊した。  
( ) 湾岸戦争が勃発した。  
( ) コール氏がドイツ首相に就任した。

## (2) 論文体テスト

「～について論ぜよ」の発問形式を採ることが多く、回答者は思うままの自由記述で回答できるという特徴がある反面、意図しないような回答が発生することがあり、また採点者によって評価が大きく変わるという問題が指摘される。理数系では、基礎から着実に学習を重ね、評価を行うことが多いので、論文体テストは割合少ない。むしろ、人文系のテストで、ある事柄に関連したストーリーを回答させたり、各自の意見をも含めて回答させるという場合に用いられる。

### 2.3.2 アンケート

ある事柄に対するイメージや意見をデータとして採取するには、一定の書式の則ったアンケートが用いられる。

#### (1) 賛否の形式

2者択一であるから、回答のしやすさの点で不特定多数にアンケートを求めるときに利用しやすい。

(例) あなたは、CAIという教育方法を授業に導入するのに賛成ですか、それとも反対ですか。

選択肢 1. 反対 2. 賛成

この形式は、回答のカテゴリーが2つしかなく、回答が困難になることも考えられるので、次のように2の中間の選択肢を用意することもある。

選択肢 1. 反対 2. どちらともいえない 3. 賛成

この形式で、ある意見を問う場合、なるべく短い文章に対して回答する形式が望ましい。もし、文章が長いと、「反対」という回答が、どの部分に対してなのかが不明確になるからである。

#### (2) 序列の形式

イメージの問う場合はそのイメージ用語をどのくらいつよく感じたか、また賛否を問う場合はどのくらい賛成(反対)かを、度合として問うのに用いる。一つの尺度の両端に対語を書いたり、それぞれの‘度合’の位置に具体的に言葉で書き表したりする。また、段数は、3, 5, 7を用いることが多い。

(例－ 5 段階の場合)

今回のテストを、どのように感じましたか。

	1	2	3	4	5
むつかしい	---	---	---	---	やさしい
ためにならない	---	---	---	---	ためになる

いずれの書式をとるにしても、アンケートは項目の文章作りや用語の選定が最重要課題である。

そこで、ここでは教育評価に限定して、評価項目などの選定に役立つと思われる文献を、以下に示しておく。いずれも日本教育工学雑誌（または日本教育工学会論文誌）の掲載論文である。ただし、どの文献も、評価そのものが中心的課題であるとは限らない。

#### I 教授スキル、講義、授業の評価に関するもの

- ・竹谷：意味構造分析の利用法と授業評価への応用、12, 1, pp. 1 - 8, 1988
- ・北尾、速水、中村：教授スキル評価の視点に関する検討、12, 3, pp.91-99, 1988
- ・吉田、佐藤：教育実習生の児童に対する知識の変化、15, 2, pp.93-99, 1991
- ・赤堀：教授・学習行動のパターン分析、13, 4, pp.139-147, 1989
- ・根本：教師の子供統制法、14, 4 pp.169-174, 1991
- ・吉田：教育実習生のリーダーシップ測定項目の作成と妥当性の検討、13, 1, pp.21-27, 1989
- ・山西、松原、山下、佐々木：3年次教育実習学生の実習自己評価に関する一考察、13, 1 pp.39-44, 1989

#### II 学習者相互交流の評価に関するもの

- ・竹谷、森本、岡田：ソシオメトリック・テストの測定解析法と活用法の一提案、12, 4, pp.139-148, 1989

#### III 教育機器や設備の評価に関するもの

- ・永岡、赤倉：教育機器・設備評価に関わる要因の分析、12, 4, pp.149-162, 1989
- ・松田、坂元：Logoを利用した小学校高学年における情報教育カリキュラムの開発とその評価、15, 1, pp. 1-13, 1991



#### IV 教材メディアの評価に関するもの

- ・赤堀、小島：刺激想起法と視聴反応分析による番組改善研究の試み、14, 4, p.133-146, 1998
- ・北垣：ファジイ積分による教育システムの評価とその応用、13, 4, pp.149-158, 1989
- ・川淵：映像資料データベース管理のための画像分類支援システム、14, 4, pp.159-167, 1991
- ・北垣、多賀谷：CAIのコースウェアの設計基準と評価概念、12, 1, pp.21-28, 1988

#### V 情意の評価に関するもの

- ・渡辺、吉崎：授業における児童の認知・情意過程の自己報告に関する研究、15, 2, pp.73-83, 1991
- ・奥野、中嶋、宮阪、上田：ホログラフィック・ディスプレイの製図学習への応用、11, 1, pp.35-44, 1987
- ・平田：高校生のコンピュータ不安を予測する要因、15, 3, pp.125-135, 1991
- ・永岡、赤倉：学習者の興味特性とキーボード入力速度のコンピュータテストとの関連、13, 3, pp.115-123, 1991

#### VI 学力の評価に関するもの

- ・北垣：ゲーム論的出題様式における確信度の算出法、15, 2, pp.101-103, 12991
- ・宮地：コンピュータを演習問題の個別化と解答照合に用いた工学演習の有効性、14, 1, pp.13-27, 1990

#### VII 経済面の評価に関するもの

- ・牟田、坂尻、坂元：コンピュータ教育の費用効果分析、14, 2, pp.61-71, 1990

#### VIII そのほか

(校内生活、心身)

- ・松原、高尾：情報化社会の中での入試制度の変化が大学新入生に及ぼした影響、145, 3, pp.131-137, 1990

(創造性)

- ・坂元、波多野、坂元：子どものコンピュータ使用と心理学的変数との関連性、15, 4, pp.143 - 155, 1992

(評価法)

- ・藤森、繁樹：大学における評価と教授法に対する学生の意識調査, 14, 2, pp.97 - 103, 1990

## 2.4 コンピュータによる評価

### 2.4.1 テストの評価

学習者の理解度をコンピュータで測るときは、多肢選択法を用いることが多い。評価の信頼性、コンピュータの記号処理能力、回答の容易さなどからこの方法が用いられているわけであるが、特に理解度の過程を追跡することを意図するときは、テスト問題を個別に評価せず、いくつかのテスト問題に関連をもたせておき、複数個の回答内容を相互に比較して評価・診断するという方法がとられる。

簡単な例をあげれば、図2.1(a)の論理回路図で所定の入力に対してある学習者の出力の回答が誤答であったとしよう。このとき、図(b), (c), (d)のそれぞれの素子の作動を理解しているか否かがわかれば、(a)でのつまづきの原因がより明確になる。図2.2はテスト問題の関連を矢印で表現したもので、このように、一つの教材を細かな学習単位にわけてそれらの関連を明確にすることによって、実際に有用な診断機能をもつテスト問題群をコンピュータ移植できることになる。

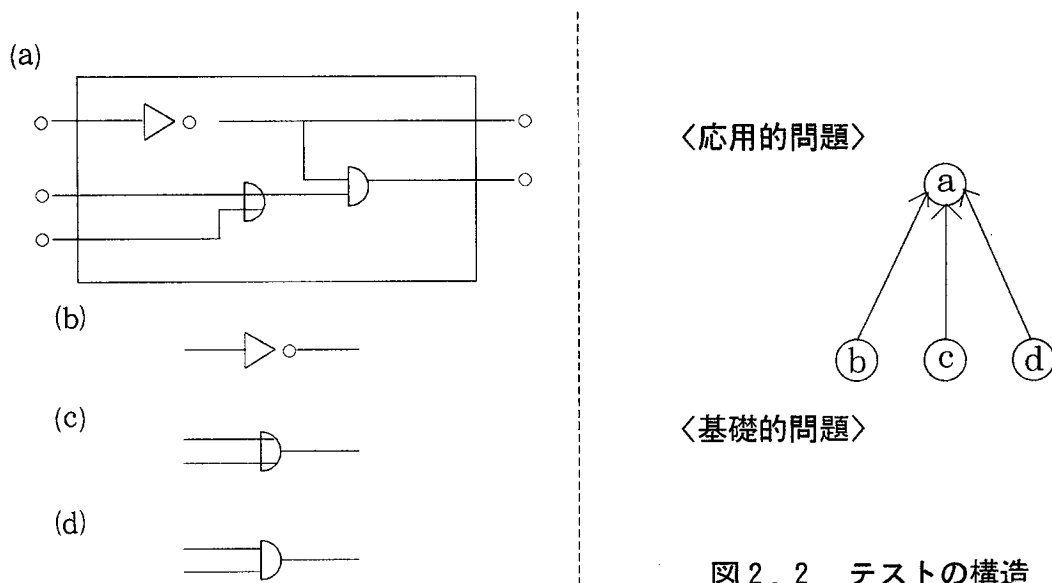


図2.1 論理回路

図2.2 テストの構造

## 2.4.2 アンケート評価

先の2.3.1でのべた序列尺度では、それぞれの数値は大小関係のみ意味をもつというのが基本的な考え方であるが、実際には間隔尺度とみなすことが多い。これは、因子分析などの多変量解析を行うためであり、多数の被験者の回答データからデータ構造を特定できるという特長がある。

実際、多数の項目を列挙しても、内容がほとんど同じとみなされるものから、存在価値の高い項目までさまざまふくまれていることが多く、項目に対する単純な平均化処理では妥当な解釈は保障されない。一定の書式に則った多数のアンケート回答は、とりあえずは、平均値、標準偏差といった基礎統計量を算出するとしても、さらに多変量解析によって、項目の妥当性などを検証したり、当該評価に関与する重要な概念用語を抽出することが必要であろう。また、多変量解析はコンピュータの処理機能が最も得意とするところである。

## 3. コンピュータテストによる確信度の評価～普及用のソフトの試作～

### 3.1 ゲーム論的出題様式

多肢選択式テスト問題で、ある一つの選択肢を回答したとき、それが高い確信度かそれとも低い確信度かの区別は難しい。しかし、次に紹介するゲーム論的出題様式(注1)を用いると、その高低を区別することが容易になる。そして、この情報は、CAIで最適な学習コースを決めるときの一つの情報になり得る。

ゲーム論的出題方式では、図3.1に示す手順で回答を行う。まず、コンピュータ画面上に設問文とあらかじめ用意した選択肢の中の一部だけを提示する。そして、未表示分の選択肢の存否を学習者に知らせずに回答を要求する。その際、表示選択肢の中からどれか一つを選べば、それを選抜的回答とみなし、正答なら報酬Pを与える。また、表示選択肢

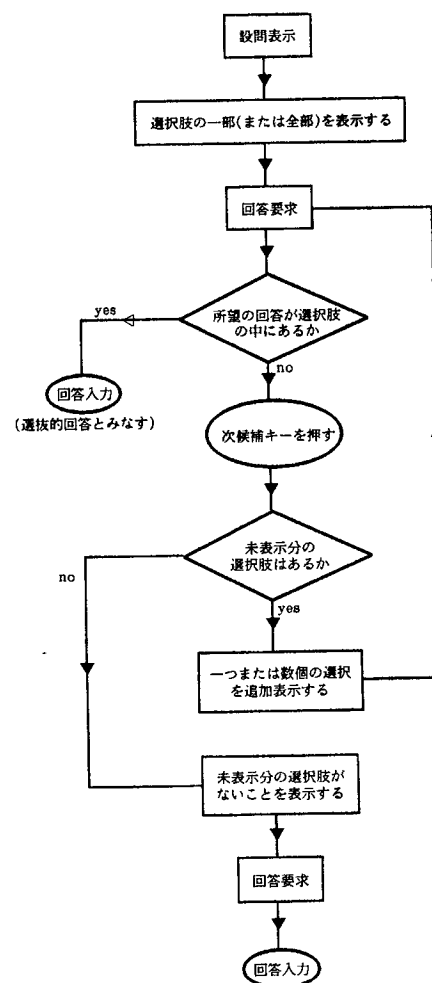


図3.1 ゲーム論的出題様式における回答入力の手順

の中に正答がないと判断した場合は、‘次候補キー’を押して新たな選択肢を追加表示させ、再び正答の存否を考えるよう、学習者に指示する。未表示分の選択肢が存在しないのに‘次候補キー’を押したら、はじめてその時点で未表示分がないことを知らせるとともに、全選択肢の中からどれか一つを選択させる。ただし、どれを選んでも消去法的回答とみなし、正答であるときは、 $Q < P$ となるような報酬 $Q$ を与える。そして、学習者にはあらかじめ報酬の額を覚えておき、なるべく高い報酬をとるように指示しておく。

回答行動は、 $P$ と $Q$ の相対的關係によって定まるので、実際には $P$ をある値に固定しておく。すると、適当な $Q$ を与えれば、同じ回答でも、確信度が高いときには選抜的に選び、確信度が低いときには消去法的に選ぶので、その回答行動から難度の高低を識別できるわけである。

### 3. 2 普及用ソフトの試作

教育訓練の担当者がこの新しい方式に興味をもたれ具体的に理解されるように、ゲーム論的出題様式を説明するCAIを開発した。その内容構成は以下の通りである。

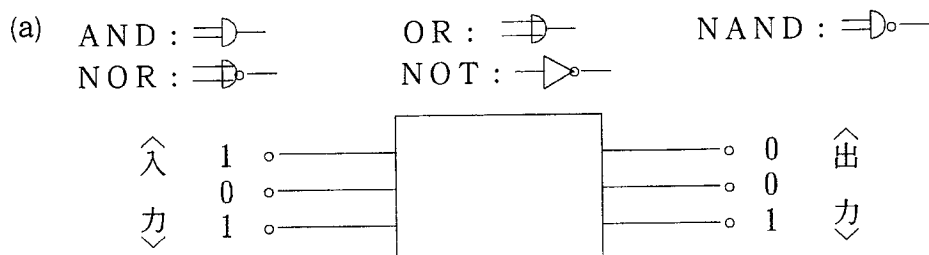
#### (1) CAIとゲーム論的出題様式

通常のCAIの方式を簡単に説明したあと、この出題様式の原理を中学程度の簡単な数学の事例で説明する。

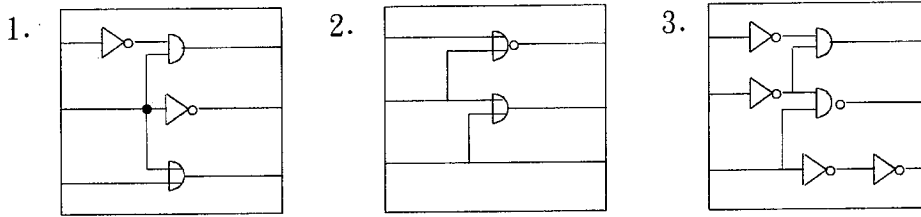
#### (2) 出題事例～論理回路～

制御関係の例として、「論理回路」を選んだ。その入力と出力を与えてその中間部分をいわばブラックボックスとして、AND、ORなどの論理素子を組み合わせ回路をその選択肢としていくつか用意しておき、対応する入出力を実現できる回路を一つだけ選択させる、という問題である。図3.2(a)がその例である。ここでは、設問数は4問とした。このソフトでは、3.1にのべた選抜的正答と消去法的正答に対する報酬は、それぞれ10と7とした。そして、利用者にはその報酬の額をあらかじめしらせておき、4つの設問の得点を累積し表示するようにした。(図3.3(a))

図に示したような入力と出力に対応するような回路図が、選択肢の中に一つだけあります。それを選びなさい。



選択肢



(b)

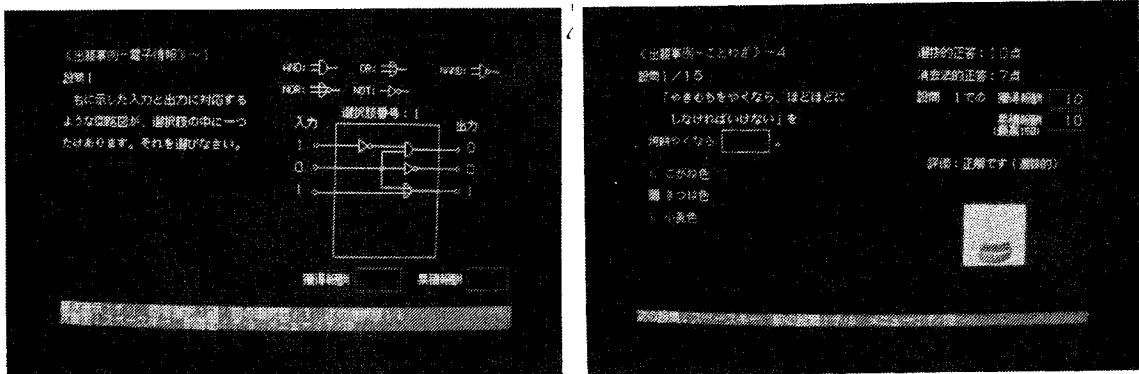
「やきもちをやくなら、ほどほどにしなければいけない」を、焼餅やくなら、…………。

選択肢：黄金色、きつね色、小麦色

図 3. 2 「論理回路」のテスト問題例

(3) 出題事例～ことわざ～

図 3. 2 (b)に示すようにことわざの設問を15問用意した。先の論理回路は、特定の人しか理解できないので、多くの人に理解できる「ことわざ」を選んだわけである。したがって、「論理回路」より「ことわざ」を量的に重視している。また、なるべくたのしい学習となるよう、各設問にはすべてイラストを入れている。(図 3. 3 (b))



(a)「論理回路」 (b)「ことわざ」

図 3. 3 コンピュータ画面 (a)「論理回路」(b)「ことわざ」

(4) ゲーム論的出題様式の要件

この出題様式はすべてのテスト問題に適用できるというわけではない。題意から肢数が特定できないこと、本来的に選択肢相互を比較せずに正答を特定できる性質のテスト問題であることなど、いくつかの要件(注1)があるが、それらを個別に説明している。

なお、このCAIソフトは、当大学で行っているCAIに関する短期実践研修の受講者6名に教材として提供した。表 3. 1 (a)は、受講者へのアンケートの内容である。項目1～4は理解面、項目5～6は情意面、項目7～8は操作面に関するものである。項目1～6は5段階評価であるが、項目7と8は発問内容の都合で3選

択肢としている。集計結果は、同表(c)に掲げた通りである。これより、項目1～6の中では、項目6が4.8と最も高い。また、最も低いのは項目4であり、ゲーム論的出題様式の要件を実際のテスト問題で具体的に示すなどすれば、この評価点がより高まることが期待される。

なお、理解面と情意面の総合評価点は、次節でのべるアンケート用コンピュータ処理システムを用いた結果、理解面が10点満点に対して約7.2また情意面が同じく8.1～8.4程度と算出されており、それなりに高得点であることが窺える（注2）。

注1：調査結果報告書第55号

注2：表3.1内の主観的重要度はこの評価点の算出に寄与しているが、主観的重要度の与え方は次節で述べる。

表3.1 開発ソフトの評価

(a)発問項目	(b)SD法による評価（段数）	(c)平均値 (標準偏差)	(d)主観的重要度	
			理解面	情意面
1. このソフトで説明している出題様式のねらいは、よくわかりましたか	全然わからなかった-よくわかった(5)	4.6 (0.48)	5	1
2. 全体的にみて、説明内容はわかりやすいですか。	大変わかりにくい-大変わかりやすい(5)	4.0 (0.00)	5	1
3. この出題様式での回答手順や回答する際の考え方はよくわかりましたか	全然わからなかった-よくわかった(5)	4.3 (0.76)	5	1
4. この出題様式で提示可能なテストの種類は、よくわかりましたか。	全然わからなかった-よくわかった(5)	3.6 (0.76)	5	1
5. この出題様式の事例としての「ことわざ」は、内容的に興味をわきましたか。	全くわかなかった-つよくわいた(5)	4.1 (0.88)	0	5
6. ‘ゲーム論的出題様式’を皆さんにとってのCAIの学習の一環としてみたとき、内容的につまらないと感じましたか。	全くつまらない-大変面白い(5)	4.8 (0.36)	0	5
7. このソフトを学習した印象として、倦怠感を感じましたか。	強く感じた-全く感じなかった(3)	3.0 (0.00)	-	-
8. 学習通行上、いくつかのキーボードを使いましたが、操作を面倒に感じましたか。	強く感じた-全く感じなかった(3)	2.2 (0.37)	-	-

## 4. 序列尺度による評価データのコンピュータ処理システム

### 4.1 ファジィ評価システム

教育訓練の評価や消費者の選好評価などを客観的に行おうとするとき、名義尺度か序列尺度にもとづいたアンケートによることが多い。このような意識調査は、明確な意見の把握と言うより漠然としたイメージの採取とみる方が妥当であり、したがって、アンケートの項目を多面的に構成し、多数のデータを統計的に処理して、考察を行う。イメージに関するアンケートは、簡単で多数のイメージ項目を用意し、評価対象に対してそれぞれの用語をどの程度強く感じるかを序列として回答するようにする。そして、多人数による回答データを収集・処理して全体的な特徴をつかむ。

このような評価データの解析では、基礎統計や多変量解析がよく用いられる。これらの処理では市販のプログラムを使用できるが、本研究では、ファジィ評価システムという、一つの新しい評価手法を開発し、コンピュータシステムを構築している(注)。

これは、ある一つの評価概念にかかわる多数のアンケート項目を序列尺度として構成し、それを多数の被験者に提示して回答してもらい、その回答データを情報集約して回答者ごとに一つの総合評価点を求めるというシステムである。ただし、唯一つの総合評価点しかコンピュータ出力されないのでは、データ量が貧弱で評価結果の解釈が困難になることがある。そこで、このシステムを用いるときは、アンケート項目を意味のあるいくつかのグループに分割し、グループごとに総合評価点を算出すると便利である。その際、

- 1) すべてのアンケート項目を多変量解析を用いてグループわけする。
- 2) すべての項目に何らかの外的基準を与えて内容別にグループわけする。

のいずれかの方法をとる。そして、グループごとに総合評価点を求め、項目の内容と照合しながら評価を進めるという手順を踏む。

### 4.2 データ処理の概要

前節の表3.1で例示したような序列尺度によるアンケート回答データに対して、総合評価点が出力される。全項目において、ファジィ評価に寄与する項目と寄与しない項目、理解面に関する項目と情意面に関する項目などを、アンケート制作者が定めているから、4.1の区分では2)に相当する。

図4.1はそのデータ入出力の概要であるが、今年度は、回答データを入力する際の操作性を改善し、より使いやすいものとした。そこで、本資料では、入力データを中心にもう一度まとめておく。

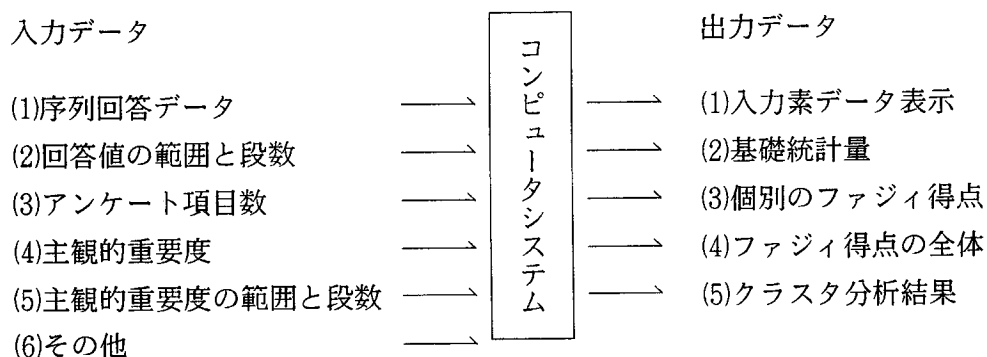


図4.1 データ出力の概要

#### (1) 各イメージ項目に対する序列回答データ

これは、図4.2(または表3.1)に示したようなアンケート項目に対する回答データである。このシステムでは、同図のように、図4.2の回答要領に指示したように評価値の中のどこか一箇所にチェックする場合と、図4.3のように、回答値を比較的高く見積った場合(上界回答値)と低く見積った場合(下界回答値)の2か所をチェックする場合の2通りを受けつける。入力ミスを防ぐため、上界回答値が下界回答値よりも低いデータに対しては、入力チェックするようにしている。

今回のコンピュータ利用学習が、これからもしばらく続くものと仮定してください。このとき、あなたご自身の行動や意思内容に関してどのような変化があると思うか、また学習環境をどのように感じるか、をおたずねします。

各項目について、それが「起こり得るかどうか」を想像し、その可能性を0～6のどこか1箇所に○をつけてください。

可能性が、

0. ほとんどないと思う。
1. ないとはいえないと思う
2. 少しあると思う。
3. ある程度あると思う。
4. 少なからずあると思う。
5. 高いと思う。
6. 非常に高いと思う。

発問項目

- a. この学習方式には、いずれ飽きる。
- b. この学習方式では、孤独感を感じずる。
- c. コンピュータいたずらや落書きをする。

図4.2 序列尺度による評価方法 I



今回のコンピュータ利用学習が、これからもしばらく続くものと仮定してください。こきとき、あなたご自身の行動や意思内容に関してどのような変化があると思うか、また学習環境をどのように感じるか、をおたずねします。

各項目について、それが「起こり得るかどうか」を想像し、その可能性を比較的低く見積った場合と高く見積った場合の2通りを考えて、0～6の2箇所(または1箇所に◎)をつけてください。

可能性が、

0. ほとんどないと思う。
1. ないとはいえないと思う。
2. 少しあると思う。
3. ある程度あると思う。
4. 少なからずあると思う。
5. 高いと思う。
6. 非常に高いと思う。

発問項目

- a. この学習方式には、いずれ飽きる。
- b. この学習方式では、孤独感を感じる。
- c. コンピュータいたずらや落書きをする。

#### 図 4. 3 序列尺度による評価方法 II

#### (2) アンケート項目数

正の整数値である。

#### (3) 回答値の範囲と段数

いずれも、整数はすべて受けつける。したがって、回答値は、図 4. 2 のように 0～6 であっても、また -3～+3 であってもかまわず、その上界値と下界値をパラメータとして入力しておけばよい。

表 3. 1 の例では、項目 1～6 に対して、その範囲を 1～5 としている。

#### (4) 主観的重要度

当該評価概念に対して各項目がどの程度重要であるかの値を、項目を相互に比較しながらデータとして入力する。たとえば、項目 B が項目 A の 1 / 5 程度の重要度しかないと思ったら、項目 A の主観的重要度に 5 を割り当てたら、項目 B には 1 を割り当てればよい。

このシステムでは、主観的重要度は、アンケート設計者が割り当てる場合と、回答者が個別に割り当てる場合の 2 通りを許している。アンケートの内容や採取状況に応じていずれかを判断すればよい。後者の場合は、図 4. 4 のように回答を求め

ることになる。

前節の開発ソフトの評価の際には、主観的重要度は表 3.1 (d)のように与えている。

コンピュータ利用学習の有用性、将来性を評価する場合、あなたは、つぎの発問項目のそれぞれを、どの程度重要と感じますか。0～4の5段階で評価してください。

0. まったく重要ではない。
1. あまり重要ではない。
2. どちらともいえない。
3. ある程度重要である。
4. この上なく重要である。

発問項目

- a. この学習方式には、いずれ飽きる。
- b. この学習方式では、孤独感を感じず。
- c. コンピュータにいたずらや落書きをする。

#### 図 4. 4 主観的重要度についてのアンケート

##### (5) 主観的重要度の範囲と段数

この範囲と段数はいずれも、先の(3)の回答値と同様、すべての整数を受けつける。したがって、主観的重要度は、図 4. 4 のように 0～4 であっても、-2～+2 であってもかまわない。その上界値と下界値をパラメータとして入力しておけばよい。

そのほか、回答者数は、1000以下で任意である。ただし、データエンドマークとして、最後の回答データの次に、任意の負数を追加しなければならない。

一方、出力データに関しては、

- (1) 入力データ —— 素データの確認のため画面表示
- (2) 基礎統計量 —— 各回答番号に対する回答頻度分布
- (3) 個別のファジィ得点 —— 回答者ごとのファジィ評価点の計算結果の表示
- (4) ファジィ得点の全体 —— (3)での計算結果の全体の分布の表示
- (5) クラスタ分析結果 —— 項目グループごとの重要度の計算結果の表示である。<sup>(2)</sup>

以上、主な入力データをのべたが、今年度の研究では、入力データのミスを極力防いだり、諸データをコンピュータ画面上に再表示し確認する際に、その内容をより詳細に同じ画面内で説明するなど、ユーザーとコンピュータの間の‘ソフト的インターフェイス’を改良を図っている。

注：操作手順の詳細は調査研究資料92号を参照のこと。

#### 文献

- (1) 橋本重治：新・教育評価法総説、金子書房、昭51
- (2) 北垣郁雄：ファジィ評価システムの開発研究、電子情報通信学会論文誌、J74-D-I, 2, pp.101-108