

## 能力領域 P「計画すること」(Planning)

2 点

♥単純な手順を計画している。

4 点

♥根拠の確実な証拠が得られるように計画している。

♥準備物やその他の証拠の出所をリストにしている。

6 点

♥自分の課題に対して科学的知識と理解を用いた計画を立てる。

♥どのような事柄が、いかに調査がうまく進むことに影響を与えるか、そして自分の計画がいかにそれらの事柄を変化させたり抑制（コントロール）させたりするか、について述べている。

♥なぜそれらの事柄を検討するかが重要であるかについての科学的な理由付けをしている。

♥自分が起こると考えていること、及び科学的理由について述べている。

♥自分がどのような証拠を得ようとしているか、及びどのくらいの量の証拠を必要としているかについて述べている。

8 点

♥自分が実行するために計画しようとしていることについて詳細に記述している。

♥自分の考えている計画がなぜその課題を実行するのに一つの良いやり方であるかの説明に、詳細な科学的理由付けを行っている。

♥自分が得る結果が正しくできるだけ正確であるものとなることを確かなものとするために、どのように準備物を用いるかについて説明している。

♥自分が起こると考えていることを述べ、またこれの説明に詳細な科学的理由を与えている。

♥自分の計画を支持する何らかの過去の研究を記述している。

♥自分の計画を支持するために、書籍や CD-ROM、インターネット、その他の出所から得られた何らかの情報を与えている。

## 能力領域 O「証拠を得ること」(Obtaining Evidence)

2 点

♥ある安全なやり方でいくらか証拠を収集している

4 点

♥課題を実行する際に、自分が後で見出したことを述べるができるように、十分な証拠を確保している。

♥自分の結果を記録している。

6 点

♥できるだけ正確に証拠が得られるように準備物を用いている

♥自分の証拠が十分な範囲に渡るようにしている。

♥自分がある結論を導くことができるように十分な証拠を確保するようにしている。

♥もし自分の証拠がかなりばらついていると思えば、可能な場合に何度か繰り返して測定を行っている。

♥自分の証拠を正確に記録するようなある明確なやり方を用いている。

♥明確な表題と正しい単位をもつ結果の表を用いることを考慮している。

8 点

♥正確な証拠を得るのに役立つ準備物を用いている。

♥平均的な測定値を得るために結果を繰り返して得ている。自分の証拠が信頼できるものであることを確かめている。

♥ある明確で正確なやり方で証拠を記録している。

能力領域 A 「証拠を分析し考察すること」 (Analysing and considering evidence)

2 点

♥自分の証拠から見出したことを述べている。

4 点

♥自分の証拠の中によりはっきりと何らかのパターンが示せるようなやり方を選んでいる。

♥円グラフや棒グラフ、図表、あるいは自分の証拠をよりはっきりと示すやり方を用いている。

♥自分の証拠の中にどんなパターンや傾向が見られるかを述べている。

6 点

♥例えば、図表や図解、線グラフを用いたり、自分のデータをうまく利用できるように役立つ計算を行うなどして、自分の証拠をはっきりと示す最良の方法を用いている。

♥もっとも適切に当てはまる線になっているか？

♥自分の証拠を用いるとともにそれに何らかの処理を施して、ある道理にかなう結論を導き、何が見つかっただかを説明している。

♥結論の中で、証拠を用いて、科学的な説明を行っている。

8 点

♥例えば、図解やグラフ、計算など、自分の証拠を処理する最良の方法を用いている。

♥この処理によって、調査に対するある意味のある結論を導いている。

♥科学的知識を用いて、ある詳細なやり方で自分が記述した結論を説明している。

♥もし、自分が起こるだろうと考えた予測をしていたなら、自分の結果が期待したような結果になったかどうかについて述べている。

♥もし、証拠が期待したような結果であったなら、その証拠が自分の予測にどのくらいよく当てはまるかについて説明している。

♥もし、証拠が期待したような結果でなかったなら、なぜその証拠が自分の予測を支持しなかったのかについて説明している。

能力領域 E 「評価すること」 (Evaluating)

2 点

♥自分が何をしたりどんな証拠を得たかに基づいて（自分が述べてきた理由も含め）、自分がその課題がうまく調査できたと思うかどうかを述べている。

4 点

♥自分のグラフに言及して、証拠がその課題に対して十分に正確であると思うかどうかを述べている。

♥何か変則的な結果が無かったであろうか？もしあったなら、グラフのどこにそれが示している。もし無くても、自分のグラフの形状について何かを述べている。

♥より正確な証拠を得ようとするために可能性のある方法の改善について少なくとも一つ提案している。

6 点

♥自分の方法が、信頼できる証拠を与え、いつも正しいと期待できるかどうかについて述べ、自分が述べたことに詳細な理由を与えている。

♥主なパターンに当てはまらないように見えた何らかの結果を指摘して、なぜそうした違いが起こったのかを説明している。

♥ある結論を導くのに十分な証拠を得ていると思うかどうかについて述べ、自分が述べたことに詳細な理由を与えている。

♥自分の方法と証拠について検討している。自分の結論を支持するより多くの証拠を得るために、いかに自分の方法を改善できるだろうか？

図 5 生徒に対する「科学」のコースワークの評価基準に関する手引き (Edexcel (2001) "Coursework guide Edexcel GCSE in Science", UG009871, p.10 より作成。)

QCA のインタビューでは、コースワークの重要性が、繰り返し語られている。

QCA インタビュー記録：[コースワークの意義について]「コースワークの質を高めることが、14 歳以降の科学を改善するかという、答は明らかにイエスであり、科学に必要なすべての要素がコースワークを含めることによって可能となる。」

[コースワークによる評価の大切さ]「・・・わが国のほとんどの理科教師は実技を行った結果が試験に組み込まれていることを良いことだと受け止めている。・・・実験を考えたり、結果を振り返ったり、それを表現するといった探究活動の全般的な側面は筆記試験では評価できないものである。」

[コースワークに伴う困難]「・・・20 パーセントという割合はそれほど大きいものではないが、必ずやらなくてはいけないということがポイントだ。大変成績の良い生徒は、コースワークをやらなくても、最高 80 パーセントまでの得点を得ることができる。そうすると、コースワークをやらなくても A のグレードを得ることができるという計算になるが、そういう事態にならないように、コースワークを行わなければ、受験資格を与えないようにした。」

コースワークの指導の状況は、学校や教師によって異なっていると思われるが、中学校のインタビュー記録から、大まかな様子について以下のように推定することができる。

第 10・11 学年の生徒は、2 週間に 11 時間の科学の授業を受け、その多くが *Double Award Science* のコースである。生徒の多くは 2 年間でおよそ 3 つの *GCSE* コースワークをそれぞれ約 3 週間費やして作成し提出する。彼らは科学以外の教科でもコースワークを提出しなくてはならないため、それ以上の時間を費やすことはできない。その期間、教師は、実験計画、実験実施、結果の分析と考察、評価の *POAE* を各 2 時間程度かけて指導するほか、時間が足りないため、*AE* については生徒が自宅に持ち帰って作業することもある。

指導の過程で、教師は「このように書きなさい」といった具体的な指示をすることはできないが、より高い得点が得られるために改善を要する点を指摘したり、そのために必要な情報収集について助言したりすることは認められている。それによって、コースワークでは到達度の低い生徒でも 8 割程度の得点が取れるようになることが期待されている。

A レベルのコースワークとは異なり、*GCSE* のコースワークではユニークさは求められない。生徒が異なる教科で提出しなくてはならないコースワークの数は非常に多く、時間的制約も厳しい。また教師にとってもユニークなコースワークの採点は困難である。結果的に、多くの生徒の *GCSE* コースワークは、お決まりのテーマとなっている。

コースワークのテーマがお決まりのものになってきている現状について、QCA でのインタビューではかなり否定的に評価されていた。

QCA インタビュー記録：[コースワークに関する想定外の状況]「資格授与機構側では、科学の *GCSE* や A レベルの科学のコースワークは、期待された程度には効果を上げていないと考えている。というのは、教師たちが毎年同じ実験をコースワークに取り上げるような学校が 80%にもなると言われており、これは想定された状況とはまったく異なっている。・・・教師は、プラクティカルな学習が同じことの繰り返しであるべきでないとしても、コースワークで

生徒がよりよい点を取るためのソリューションを数年のうちに見つけ、それを繰り返すようになる。個人的は、コースワークを試験の要件からはずしてもよいと思う。・・・」

QCA インタビューにおいて「1つか2つのプロジェクトですべての教科を含むものであれば、より優れたスキルの評価になるのではという議論がある」と話されていることから、GCSE コースワークの指導や実施形態については、今後さらなる改革が行われるものと考えられる。

コースワークにおける「科学的探究能力」の評価は、基本的に学校内で各科学教師によって行われた後、内部と外部で一部についてのチェックを行い、問題がなければその評価結果が採用される。したがって、それぞれの科学教師が、生徒のコースワークを適切に評価できることが必要となっている。

図6は、コースワークを評価する際に用いる評価基準であり、この表記に照合しつつ、各コースワークを採点するものとなっている。図6には、2, 4, 6, 8点(E「評価する」は最大6点まで)についてのみ表記されているが、これらの表記を手がかりに、それに部分的に達していたり、達していなかったりすることで、それらの中間の点数を付ける。まったくその領域の能力を発揮できない場合には0点もあり得る。前述の教科書“Nelson Modular Science”では、生徒に対して、2点がGCSE資格の階級Gに、4点が階級Eに、6点が階級Cに、そして8点が階級Aの水準におよそ相当すると解説されている。

本章の末尾に、生徒のGCSE コースワークの事例、及び参考としてAレベルのコースワークの事例を掲載している。翻訳に当たっては、文章表現に問題がある点も含めて、できるだけ生徒の作成したままの表現を再現するよう努めたものである。GCSE コースワークについては、30満点中30点、Aレベルのコースワークについては45満点中44点という、高い評価を受けた事例である。その他、訪問時には、評価の低いコースワーク事例も多く見ることができたことから、これらの事例が、中学校と大学準備教育段階(Sixth form)での指導の結果、期待されているコースワーク像を反映したサンプルであると見なすことができると考える。

GCSE コースワークの事例は、「オームの法則の検証」という基礎的な内容について、理論的考察、予備実験に基づく材質の選択、実験の計画、信頼できるデータの収集と分析、結果の考察、結論と調査全体の評価、といった一連の「科学的探究」を展開しており、POAEの評価基準に沿って、それぞれに満点を与えられるに足る記載をしていることが、教師の採点メモから確認できる。教師の採点メモは、他の教師や外部評価者(モデレーター)が採点の判断基準をチェックする際に重要な証拠となるものである。

能力領域 P「計画すること」(Planning)

- 2点 P.2a 単純な手順を説明している。
- 4点 P.4a 根拠を確実にするような証拠を収集するように計画している。  
P.4b 証拠に関して相応しい準備物や情報源の利用を計画している。
- 6点 P.6a ある手続きを計画し伝えるために科学的知識と理解を用い、重要な諸要因を特定したり、変化させたり、抑制したり、考慮に入れたり、また、適切な場合に予測を行っている。  
P.6b 証拠を収集するために、ある適当な範囲と大きさを決定している。
- 8点 P.8a ある適切な手法を計画し伝えるために詳細な科学的知識と理解を用いており、そこに、正確で信頼できる証拠を生成する必要性と、予測をした場合に予測を正当化する必要性を考慮している。  
P.8b 計画を述べるために、適切な場合に先行研究から関連する情報を用いている。

能力領域 O「証拠を得ること」(Obtaining Evidence)

- 2点 O.2a ある単純で安全な手順を用いていくらか証拠を収集している。
- 4点 O.4a 活動に十分で適切な証拠を収集している。  
O.4b 証拠を記録している。
- 6点 O.6a 十分に体系的で正確な証拠を集め、また、適切な場面で繰り返しや確認をしている。  
O.6b 収集した証拠を明確にかつ正確に記録している。
- 8点 O.8a ある適切な範囲で信頼できる証拠を得たり記録したりするための正確な手順と技能を用いている。

能力領域 A「証拠を分析し考察すること」(Analysing and considering evidence)

- 2点 A.2a 証拠によって何が示されるかを簡潔に述べている。
- 4点 A.4a 証拠を説明するための基礎として、単純な図解や図表やグラフを用いている。  
A.4b 証拠中の傾向とパターンを特定している。
- 6点 A.6a ある結論に向けて証拠を処理するために、相応しい図解や図表、グラフ（適切な場合に最適に当てはまる線が引かれている）、あるいは数字で表す方法を作ったり用いたりしている。  
A.6b 証拠に合致するある結論を導き、それを科学的知識と理解を用いて説明している。
- 8点 A.8a 詳細な科学的知識と理解を用いて、証拠を処理して導かれた根拠の確かな結論を説明している。  
A.8b 予測がなされていた場合に、どの程度まで結論がその予測を支持するかを説明している。

能力領域 E「評価すること」(Evaluating)

- 2点 E.2a 用いた手順や得られた証拠に関連した批評を行っている。
- 4点 E.4a 何らかの変則を特定しながら、証拠の質について批評している。  
E.4b 手順の適切性について批評し、また適切な場合に、それを改善するための変更点を示唆している。
- 6点 E.6a 証拠の信頼性と、それが結論を支持するに十分かどうかについて、変則を説明しながら、批評的に考察している。  
E.6b 付加的に関連する証拠を与えるさらなる研究について、詳細に記述している。

図6 「科学」のコースワークの評価基準 (Edexcel (2000) "Specification Edexcel GCSE in Science: Single & Double Award A", UG008983, pp.66-69より作成。内容は、他の2つの資格授与機構に同一である。OCR(Oxford Cambridge and RSA Examinations) (2001) "COURSEWORK GUIDANCE BOOKLET GCSE Science", Assessment and Qualifications Alliance (AQA) (2003) "AQA GCSE 3462") この配点の記述は階層的に設計されている。

⑥ 「コースワーク」の評価結果の信頼性を高める工夫をしている。

コースワークは、基本的に学校の科学教師が採点するため、共通の評価基準を元に採点したとしても、教師間で、評価基準の解釈に大なり小なりのずれが生じる。このずれが大きい場合には、認定資格の信頼性が低下してしまうため、いかにこの評価の信頼性を高めるかは重要な課題となる。

一つは、学校内での評価のずれを調整(モデレーション)すること、そしてもう一つは、資格授与機構側でいわゆるモデレーターによる外部審査によって評価のずれを調整することの2つの取り組みが行われている。

中学校でのインタビューにおいて、そうした対処の詳細が語られている。

中学校インタビュー記録：[モデレーターとの評価のずれ]「モデレーターは教師の採点と自分自身の採点を比較して、(60点満点中) プラスマイナス4点の範囲以内に収まっているかを調べる。それがその範囲を超えていた場合には、教師の採点を変更させる。私とスティーブがここに来てからは、このような変更が行われたことはない。・・・私とスティーブが中心になって、実例を元にどのポイントに何点を与えるかというような採点方法についての研修をかなり行った。」

[校内モデレーションを行う教員の苦労]「コースワークの採点は本当に大変で、昨年度も第11学年の生徒の作品を少なくとも750は採点しなくてはいけなかったが、一人一人の作品を全て我々が採点しているというわけではない。モデレーターがやるように優秀なもの、下位のもの、真ん中のものからそれぞれ4つのサンプルを取り出し、それがきちんとできていれば、他の採点についても大丈夫であろうということで、その教師の採点を認める。もしもそれがいうまくいっていないようであれば、全ての作品の採点を見直す。」

[モデレーターによるチェック]「モデレーターのサンプルの選び方は、成績の上位から、下位までの全体から、幾つかの部分に分けて満遍なく選ぶ。教師がどの程度の援助をしているかということは学校により差がある。初めて探究的な活動を評価しようとした年には、特に教師が問題になった。彼らが教えすぎてしまったのだ。モデレーターが入ってすべての作品をチェックした。モデレーターは実技の試験を行うようになってからずっと存在してきた。実技の試験が導入されてからずっとあったはずだから、20年以上も前からということになるだろうか。最近ではモデレーターが入ることはほとんどないのだが、インターネットからの情報をそのまま書き写して、自分の作品として提出されたときにあった。現在大変たくさんの情報がネット上に存在している。そのような不正を発見するためには、こちらもサーチエンジンなどを使ってある特定の表現やグラフを引き出し、生徒の作品と比較するというようなこともやって見なくては行けない。またこれまでの作品に比べて、文章に間違いがなかったり、つづりが間違っていないということで見つかる場合もある。」

このように、各学校の教師が多大な労力を払ってコースワークの評価の信頼性を高めており、また、評価結果を調整するモデレーターという専門職も必要不可欠な存在として確立されていることがわかる。

また、資格授与機構側の取り組みについては、Edexcel でのインタビューで次のように語られている。

*Edexcel* インタビュー記録：[コースワークの評価の客観性を高める]「実技の採点は、グレード別に採点方法を示したものがあり、すべての教師はそれを参照して採点しなくてはならない。・・・その生徒が好きであるとか嫌いであるとかの主観的な判断から離れて、客観的になることを求められる。これを「職務上の判断」と呼んでいる。これに従って教師は評価しなくてはならない。授与機構はある程度の許容範囲を持って各学校の評価を受け入れているが、その許容範囲を超えた評価が提出された場合には、授与機構から試験官を派遣して、コースワークが適正に評価されていない旨を伝える。GCSE や GCE ではすべてのコースワークを提出して、チェックを受けることになる。いくら評価方法についてのガイドを配布したとしても、評価を客観的に行うということは難しい。だから授与機構ではコースワークに用いる実技の例を示し、評価方法についての情報を提供する。また、学校の評価システムもチェックする。これらが GCSE タイプの試験について資格授与機構ができることである。教師の評価技能を高めるためのセミナーも開設している。そこでは、実際にいくつかのコースワークを評価してみて、試験官の行った評価と自分の評価との違いを知り、許容範囲内で評価が行えるように訓練される。」

[コースワークの指導過程における教師の助言の許容範囲について]「・・・授与機構は学校から出された評価について審査し、高い得点を与えすぎていると判断した場合、その評価を下げるということもある。たとえば提出された20人の生徒のコースワークがすべてAの判定であるとすると、「これはおかしい。」ということになる。一般的に成績には、正規分布どおりにならないとしても、ばらつきが存在するはずで、すべての生徒の判定がAになっていればもう少し(多くの生徒の作品を提出させて)詳しく見なければいけなくなる。成績のよい学校と、そうでない学校ではそのばらつき具合が異なっているだろうが、モデレーターは、疑わしい評価の分布を見抜く訓練をしている。」

## 第7節 生徒は意欲的に科学的探究に取り組んでいるのだろうか？

### ⑦ 学ぶ側も教える側も実践的な「科学的探究」の大切さを理解している。

インタビュー調査では、英国における生徒の科学への良好な意識の背景を探るため、生徒の科学への学習意欲に関わる質問を問いかけてみたが、それに対する複数の回答は、学ぶ側にも教える側にも実践的な「科学的探究」の大切さが理解されているという印象を強くするものであった。つまり、英国における科学や科学的探究への学習意欲は、それが好きか嫌いかといった情意的な意識よりも、それがいかに大切かというより認知的な価値判断によって支えられているのではないかと考えられるのである。

*QCA* インタビュー記録：[コースワークによる評価の大切さ]「・・・わが国のほとんどの理科教師は実技を行った結果が試験に組み込まれていることを良いことだと受け止めている。生徒もまた、その必要性を理解している。」

*Edexcel* インタビュー記録：[実際的な活動が生徒を動機付ける]「・・・科学を好きにさせるひとつの大きな要因として、多くの実習を行うことがある。理論的な側面を無視するわけにはいかないにしても、より多くの実技を取り入れることが生徒を動機付けることにつながるであろう。実技だけに限ったことではない。もっと探究的に授業をすることも大切である。ただ前に立って、知識を伝達するのではなく、生徒とのやり取りの中から授業を進めていくことが大切だ。」

中学校インタビュー記録：[生徒はコースワークの大切さを理解している]「彼らはコースワークが好きである。なぜならコースワークがどれだけ有用なものであるかわかっているからだ。・・・彼らはコースワークの大切さを理解しているし、やらなくてはならないと思っている。それに評価されるということには抵抗はあろうが、コースワークの過程は嫌いではない。コースワークが自分の能力を最大限に発揮できることも理解している。締め切りに追われるのがいやなだけだ。よくできる生徒については何の問題もない。彼らはコースワークを楽しんでやっている。問題なのは、ある程度の能力があるのに、期限を守って学習を進められない生徒たちだ。」

[実験が好きである以上のことが必要]「実験はとても好きだから、試験を受けるより、実技試験を好む。・・・特に下の学年の生徒たちは、実験をととても楽しんでいるのだが、探究の過程を意識して何のためにやっているかを意識していないことがよくある。純粋に実験という活動を楽しんでいるだけなのだ。だから実験の目的などを書くことは好きではない。どうしてこの実験をすることが必要なのか。というようなことについてかなりの時間を割いて教えるようになった。」



## まとめと考察

第2章では、まず、英国のQCA、Edexcel、小学校、及び中学校での訪問調査についてインタビューの結果を記述した。そして、インタビューの記録と、収集した資料等を元に、「科学的探究能力」の指導と評価について考察し、次の特徴を明らかにした。

- ① 系統的に「科学的探究能力」が指導されるように情報提供を工夫している。
- ② 教科書において「科学的探究能力」の指導が組み込まれている。
- ③ 日常的な科学の授業を通じて「科学的探究能力」の指導と評価を工夫している。
- ④ 全国テストで筆記試験による「科学的探究能力」の評価を工夫している。
- ⑤ 資格試験の「コースワーク」により「科学的探究能力」を指導し評価している。
- ⑥ 「コースワーク」の評価結果の信頼性を高める工夫をしている。
- ⑦ 学ぶ側も教える側も実践的な「科学的探究」の大切さを理解している。

さて、英国の科学カリキュラムに関して見出されたこれらの諸特徴から、わが国の科学カリキュラムを振り返り考察を行う。

まず、わが国においては、学習指導要領において「科学的探究能力」を系統的に指導するような設計となっていないため、どのような科学的探究能力をいつの段階でどの程度まで育成するかが明らかでない。結果として、科学カリキュラムの中で、意図的計画的に効率よくより高度な科学的探究能力が育成されているとは言えないであろう。

したがって、教科書も、科学の内容面での知識理解中心の構成となり、科学的探究能力が徐々に深められ発展していくような構成となっているとは言えないであろう。

また、わが国でも日常的な理科授業において、観察実験を通じた理科学習が強調されているところではあるが、それを通じて育成すべき「科学的探究能力」が曖昧で、ややもすると、科学的に探究すること無く、手順に従って観察実験という行為を経験したに過ぎないものとなりかねない。また、英国で強調されているような「実験計画」と「証拠の獲得」、「証拠の分析と考察」、及び「評価」まで含めた一連の「科学的探究」過程には、相応の時間が必要であり、授業時数の少ないわが国の理科授業において、そのような「科学的探究」を通じた「科学的探究能力」の指導は物理的に一般的に困難であると考えられる。したがって、そのような一連の「科学的探究」に携わる機会を工夫して設定することが、「科学的探究能力」を意図的計画的に育成するための重要な要件であると考えられる。

英国の科学カリキュラムでは、ナショナルテストや資格試験のコースワークを通じて、「科学的探究能力」の指導と評価に非常な労力と時間、情熱、知恵が注ぎ込まれていた。これは、わが国の科学カリキュラムと大きく異なる点である。わが国においても科学的探究能力の育成は一般的には重要な教育目標であるが、それが具体的な教育内容として系統的に規定されていない現状において、それが主たる指導と評価の対象とはなりにくい。一方、英国においては、科学カリキュラムが、科学の概念的体系と、科学的探究能力の発達の体系の2つの体系が車の両輪のように組み込まれて編成されているため、科学的探究能

力の育成がこのように重視されていると考えられる。ここに、両国での科学カリキュラムの編成原理に大きな違いがある。

英国では科学を学ぶ側も教える側も実践的な「科学的探究」の大切さを十分認識した環境の下で、「科学的探究能力」の育成が図られていた。わが国で、理科を学習することの大切さを意識している生徒が少ない（第1章図2）状況を改善するためには、科学カリキュラムにおいてこうした価値意識の醸成に取り組みねばならないと考える。

さて、今日わが国の高等学校では、SSH（スーパーサイエンスハイスクール）や理数科などにおいて、生徒の科学の課題研究を重視した科学カリキュラムが開発され広まりつつある。しかし、そのような実践的な科学的探究の経験を積んだ生徒の学習成果を適切に評価する手法や評価の基準については、未だ開発されているとは言い難い状況である。英国のGCSEやAレベルの資格試験におけるコースワークの評価手法は、高い妥当性と信頼性をもつ課題研究レポートの評価法として、わが国でも活用可能なのではないであろうか。その際には、あらかじめ、どのような評価基準で課題研究レポートが評価されるのかを生徒に理解させた上で、課題研究を実行させ、提出されたレポートを客観的に評価し、評価結果を指導に生かすという、これまでとは異なった教授学習過程を採ることとなる。今後、実践的に研究を行ってみる必要があると考える。

また、わが国においては、多くの小学生が、夏休み期間に探究的な観察実験を伴う科学研究を意欲的に行い、科学論文を作成するというすばらしい文化が存在している。その多くは、自ら調査課題を設定し、観察や実験を計画し、証拠を集め、分析し考察するという、一連のオーセンティックな科学的探究を論文にまとめたものとなっている。しかし、中学生、高校生となると、そうした文化は陰を潜め、大半の生徒にとって自主的な科学研究は無関係な存在となってしまふ。これを有効に活用し発展させるための一つの可能性は、小学生に始まる科学研究が、発達とともに次第に高度な科学的探究へと発展するような指導過程を科学カリキュラムに組み込むことであると考えられる。その指導を通じて、すべての子どもたちに「科学的探究」の大切さを理解させるとともに、実社会で求められる実践的な科学的探究能力を身につけさせることが可能となるのではないだろうか。

## GCSE コースワークの事例 (GCSE Physics, POAE ともに満点)

(できるだけ生徒の記述と教師の書き込みに忠実になるよう訳した)

### 予備実験:

#### 基礎理論:

電流は、導線を通して、電子によって運ばれる。電子は、イオンを介して移動し、電子は抵抗がどのくらい高いか低いかを決定する。抵抗を変化させるものには4つあり、それら4つの事柄は、温度、太さ、材質、及び長さです。しかし、それについてはここで深入りしない。金属中のイオンは、位置が固定されている。電子が導線を通るときに、電子はイオンに衝突し、それで運動エネルギーを失って速度を落とす。そうして抵抗がより高くなる。それで、材質中により多くのイオンがあることは、一般的により多くの抵抗を材質が持つことを意味する。

なぜ導線中の抵抗が異なるかは次の理由による。

#### 温度:

導線の中の温度が増すにつれて、一カ所に固定されたイオンが振動し始める。というのは、イオンの温度があがるにつれより多くのエネルギーをもつので振動できるのである。このことは電子にとっては問題となる。というのは、電子が導線を伝わって移動する際に、イオンがより動くことで、電子がより障害に衝突するからである。これは、抵抗がより高いことは、電子がイオンによって減速されるからだということの意味する。

#### 材質:

導線の材質はその抵抗に影響を与える。というのは異なる金属は多かれ少なかれイオンを持っており、もし金属がより多くのイオンを持っているなら、より少ないイオンを持っている導線に比べて、その抵抗が高くなるからである。

#### 太さ:

導線の太さはその抵抗に影響を与える。これは、もし導線がより小さい断面積のものよりもより広い断面積をもつなら、より広い断面積がより小さい抵抗となる、その理由は、電子が導線を移動するのに、細い導線よりもより多くのスペースをもつからだ。例えば、より広い導線は例えば自動車道路の一つ以上の道であって、したがって、交通がより速く進むように、抵抗はより低い。もし、より細い導線であるなら、田舎の道路で、交通がとてもゆっくりと進むように、抵抗はより高い。

#### 長さ:

長さもまた抵抗に影響を与える。これは長さが増すにつれて、電子がより多くのイオンを通過しなくてはならないからだ。これは抵抗に影響ないはずだ、というのは  $R=R_1+R_2$  の公式による。長さが抵抗に影響する理由は、より短い導線がより速く導線が暖まるからだ。これは、イオンがより速く動いて、電子が減速されて、抵抗がより高くなるということの意味する。

part

### 予備実験:

#### 目的:

目的は、どの種類の導線が単位長さ当たりで最も高い抵抗か、及び最も低い抵抗かを決定することである。

予備実験において、異なる材質を用いることが抵抗にどのくらい影響を与えるかを調べる。例えば、銅は、導線の太さが 0.5mm で長さが 100cm のときに 0.3 オームの抵抗を持つ。この情報は、コンピューターシミュレーションを用いて見つけたものだ。しかし、これはすべての金属で同じではない。公正に確かめるため、それぞれの材質の長さと同直径を同じにした。また、導線の温度も導線の抵抗に影響を与えるので、その温度を変えないようにする。

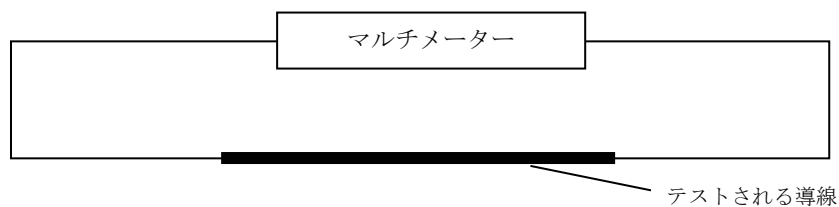
**方法：**

50cm の長さの金属線を選び、それぞれワニ口クリップを使って、マルチメーターにつないで、回路にする。マルチメーターは電源を内蔵しているので、それを電源につなげる必要はない。マルチメーターは、導線の長さに伴う抵抗を測定しようとしている。次のような 4 種類の異なる金属についてこれを繰り返す。

- 銅
- 鉄
- ニクロム
- コンスタンタン

より正確な結果を得るために、それぞれの導線について 2 回測定し、その平均値を取る。

**図解：**



**結果：**

材質	抵抗 ( )		
	1 回目	2 回目	平均
ニクロム	17.2	21.7	19.5
鉄	4.2	4.5	4.4
コンスタンタン	10.0	11.6	10.8
銅	1.1	1.0	1.1

**結論：**

結果から、抵抗が平均で 1.1 と最も小さかった銅が最も電気を通し、抵抗が平均で 19.5 だったニクロムがもっとも電気を通さないことがわかった。私は、ニクロムの計測は測定値同士がかなり離れているので、最も信頼性が良くなかったと思う。本番の実験では、最も高い抵抗値であったニクロムを用いようと思う。その理由は、長さの違いが、より大きな抵抗の違いになって現れ、より正確な結果をもたらすと思うからだ。

P8b

## 1本の導線（ニクロム）の抵抗値を見つける実験

### 目的：

この実験の目的は、長さがどのくらい導線の抵抗値に影響するかを調べることである。

### 計画：

#### オームの法則

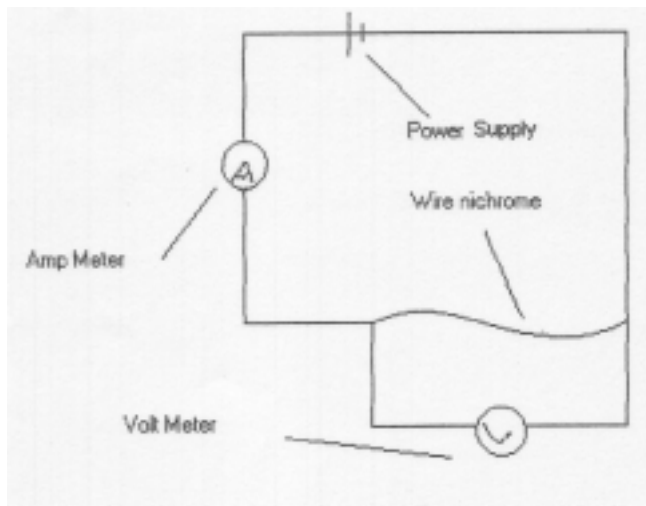
オームの法則は導線の抵抗値を計算するために用いられる。これは、実験がどのくらいの抵抗を導線が持っているかを見出すことであるので、オームの法則がとても便利である。オームの法則を用いるのは、ある決まった長さの導線の抵抗を調べなければならないからである。

P6b  
part

オームの法則の公式は次のものである。  $R = V/I$

R=抵抗 V=電圧 I=電流

### 実験



### 器具

- 電流計
- 電圧計
- 電源
- 導線
- 1 mのものさし

P4b  
P2a  
P4a

電流と電圧、導線の長さを測定する。これらすべては、正しく測定し、どれ 1 つを読み落としてしまうと正しい抵抗値が計算できないので、とても大切だ。

電流計の最大は 10 アンペアまで測る。

電圧計の最大は 20 ボルトまで測る。

### いかに実験を公正に保つか

この実験をできるだけ公正に保つために、次に上げるさまざまな事柄を行う。

- 実験を繰り返すたびに、同じ装置を用いる。これは、もし装置を実験を繰り返すたびに変わると、結果が変わってしまうからだ。これは、装置を変えると、装置自身がわずかに変化することによる。

P6b  
part

- ・ 実験を通じて、温度を同じに保たなくてはならない。これは、導線の温度が上がると、導線中のイオンの運動エネルギーが増すからである。それでイオンがより速く動き、電子がより導線中を移動するのがよりむずかしい仕事となるので抵抗が増す。これを抑制するために、測定を迅速に、かつ、電源を同時に入れる。
- ・ 実験を同じに保つために、その他のすべての要因を同じにしておかなくてはならない。

### 安全

O2a

- ・ この実験を安全に行うため、高い電圧で短い測定を行わないようにする。導線の温度が増すにつれて、危険度が増し、手をやけどするかもしれない。

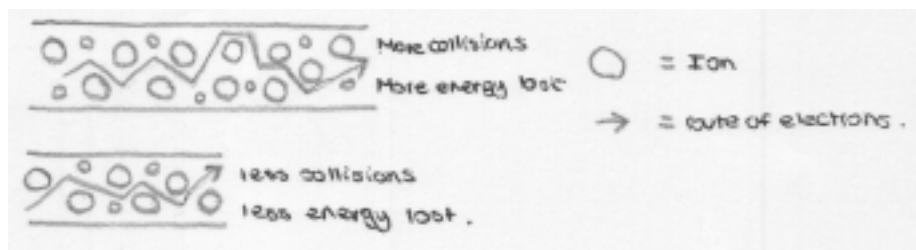
P6b

表では、1mm の精度で測定しようとする。20,30,40,50,60,70,80,90,100cm の長さで測定を行う。平均の測定を行う。

P6a  
part

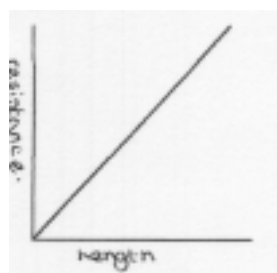
### 予測

私の予測は、導線の長さが増すにつれて導線の抵抗が同様に増すということである。これは、導線の長さが増すと、より多くのイオンと電子が衝突するからだ。これは、電子が運動エネルギーを失い、減速し、抵抗が増すということを意味する。電子が長い導線と短い導線を移動する様子を下に示す。



また私は、長さが2倍になると、 $R=r_1+r_2+r_3$  の公式から、抵抗が2倍になると思う。それは、もし20cmで2Aであれば、それを40cmとすると、もう一つの20cmになるので、もう1つの2Aになる。それで $R=2A+2A=4A$ になり、長さが2倍になると抵抗が2倍になるということの意味する。

私が期待するグラフはこのようになる。



P8a

グラフがこのようになると期待する理由は、 $R=R_1+R_2+R_3$  の公式で、もし導線の長さが2倍になると導線の抵抗が2倍になることを意味するからだ。この公式を用いるのは、回路が直列回路になっているからだ。

### 観察

	長さ(cm)	電圧			電流 (I)			抵抗		
		1	2	3	Av	1	2	3	Av	( )
	20.0	4.71	4.71	4.70	4.7	0.62	0.61	0.61	0.61	7.70
O4a	30.0	4.82	4.84	4.82	4.82	0.42	0.32	0.43	0.42	11.47
O4b	40.0	4.90	4.91	4.90	4.90	0.32	0.33	0.33	0.32	15.31
O6a	50.0	4.95	4.96	4.95	4.95	0.27	0.26	0.26	0.26	19.03
O6b	60.0	4.98	4.99	4.98	4.98	0.22	0.22	0.22	0.22	22.63
O8a	70.0	5.00	5.01	5.01	5.0	0.19	0.19	0.19	0.19	26.31
	80.0	5.03	5.03	5.02	5.02	1.17	0.16	0.17	0.17	31.37
	90.0	5.04	5.04	5.05	5.04	0.15	0.15	0.15	0.15	33.60
	100.0	5.05	5.06	5.06	5.05	0.13	0.13	0.13	0.13	38.84

### 分析

A2a

私のグラフは、長さが増すにつれて抵抗値が増している。しかし、興味深いことに、長さが2倍になると抵抗もそうなっていることをグラフは示している。私のグラフと結果は、私の予測が正しかったことを示している。これは、私が長さが2倍になると抵抗も同様に2倍になると予測したことによる。これは、 $R=R1+R2$ の公式によるもので、私の結果はこのことを結論する。

	長さ(cm)	電圧			電流 (I)			抵抗		
		1	2	3	Av	1	2	3	Av	( )
	20.0	4.71	4.71	4.70	4.70	0.62	0.61	0.61	0.61	7.70
	40.0	4.90	4.91	4.90	4.90	0.32	0.33	0.33	0.32	15.31
	80.0	5.03	5.03	5.02	5.02	1.17	0.16	0.17	0.17	31.37

A6a

公式  $R=R1+R2$  は、直列回路でのみ使える。公式は、長さが2倍になると抵抗が2倍になることを意味する。 $R=r1+r2+r3$ は、もし20cmで2Aであれば、40cmにすると、もう1つの20cmとなり、そのためもう1つの2Aとなる。

$$\text{抵抗} = \frac{\text{抵抗係数} \times \text{長さ}}{\text{断面積}} \quad R = \frac{\rho l}{A}$$

A8a

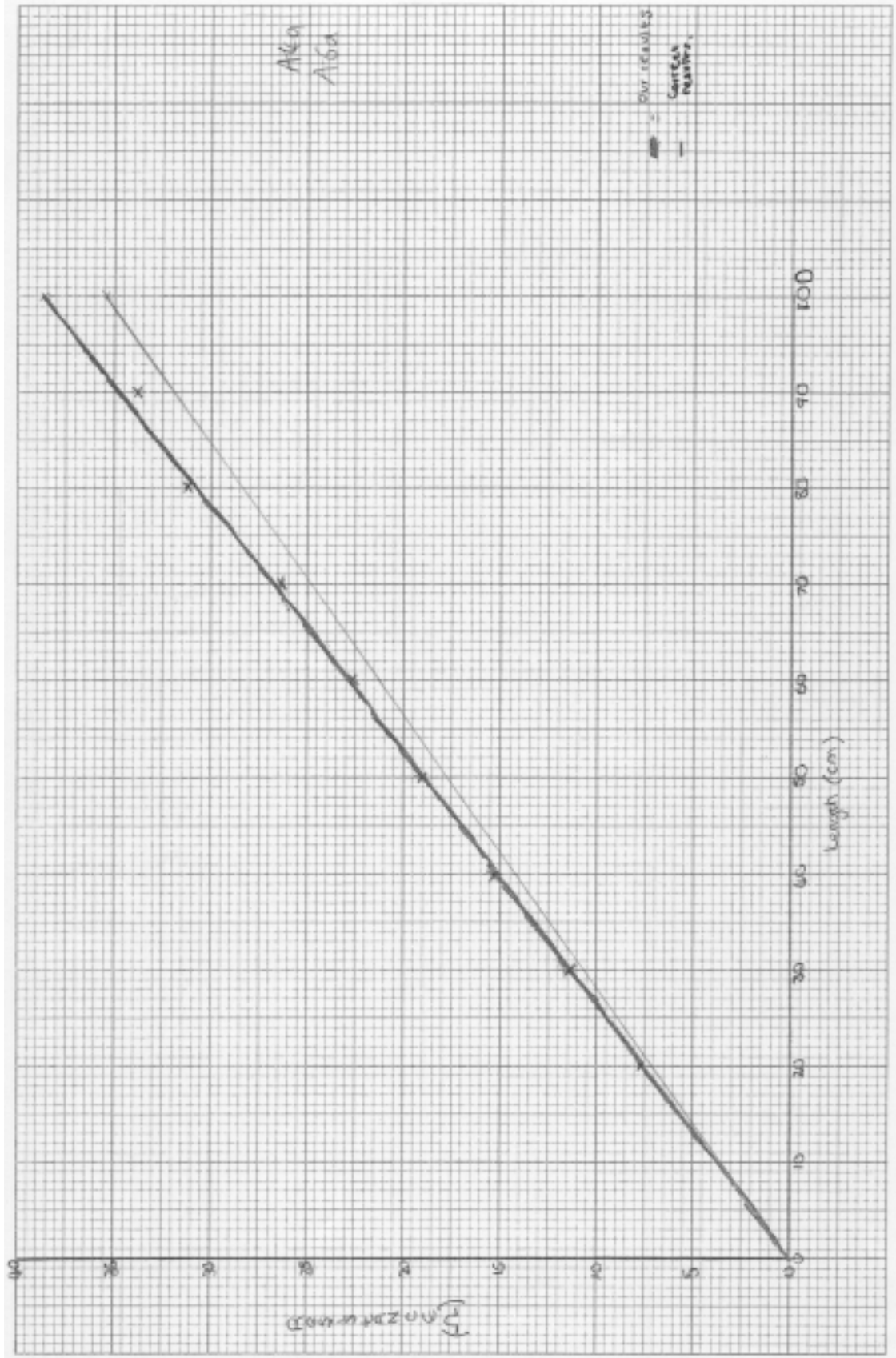
実験ではPと断面積は一定である。そのため、公式はRが直接長さ(L)に比例すると予測する。結果はこのことを示している。これは、長さが2倍になると抵抗が2倍になることを示している。ここで私は、長さが2倍になると、導線の長短によって、抵抗が2倍になることがわかる。一本の長い導線には、より大きな抵抗がある。これは、電子が導線を通ってずっと遠くまで行かなくてはならないからだ。これは、電子と衝突する道により多くのイオンがあることを意味している。電子の通る道のりが長いと、イオンとのすべての衝突のために電子の運動エネルギーがより少なくなる。短い導線では、抵抗はより小さい、というのは、電子は長い導線を伝わるように遠くまで行かなくて済むからだ。もし、短い導線を通して伝わるならば、運動エネルギーを減速する電子はそう多くない。それで、導線をより速く伝わるので、より小さい抵抗となる。長さが増すとともに電流は減少する。これは、導線が長くなると、電子が通過するのにより多くのイオンと衝突するからだ。それで長さが増すとイオンも増すので、電流が減速する。

A8b

A4b

理論的には欠陥があるが、なお、A8bを当てるに十分である

### 評価





E2a 私の結果の表とグラフは、導線の抵抗への長さの効果の証拠となる。結果の信頼性は可能な限り正確であるとは言えない。グラフ上では、重大な変則は無いが、直線からわずかにずれた結果が2つある。これは、たぶん次のような事柄によるものだろう。

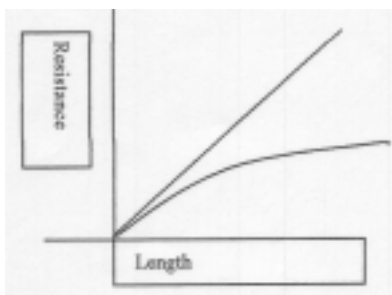
E4a

これは、たぶん、どれくらい長い時間導線に電気を流すかを測っていなかったからである。このことは、導線が温まって電子が運動エネルギーを得ると、抵抗をより小さくするという問題を引き起こすだろう。これが結果を変化させうる。

E4b もう一つ結果を変化させたかもしれないことは、ワニ口クリップは分厚かったが、もしそれがより薄かったら実験をかなり改善しただろう。ワニ口クリップが結果を変化させたかもしれない理由は、抵抗の長さを正しく測ることがそれによって難しいことがわかったからである。

グラフ上の結果は、ニクロムの真値として公表されている正しい結果とわずかに異なっている。私たちの結果で、長さが短いところの方が、グラフの終わりの当たりの長さが長いところよりも、より接近している。たぶん、同じ装置を使ってもふたたび同じ結果が得られるだろう。

もしこの実験を繰り返すので有れば、より長い長さを用いるべきだ。というのは、抵抗が長さに比例して大きくなるのか、抵抗が水平になり始めるのかがわかるからだ。



**確固とした結論：**

結果は一つの確固とした結論を支持するものである。それは、結果が抵抗が増すとともに長さも増すということを語っているからである。それで、もし長さが2倍になると、抵抗も2倍にある。下の表が示すように。

長さ(cm)	電圧			電流	電流			抵抗	
	1	2	3	Av	1	2	3	( )	
20.0	4.71	4.71	4.70	4.70	0.62	0.61	0.61	0.61	7.70
40.0	4.90	4.91	4.90	4.90	0.32	0.33	0.33	0.32	15.31
80.0	5.03	5.03	5.02	5.02	1.17	0.16	0.17	0.17	31.37

E6a この表は100%正確ではない、というのは、用いている装置が可能な限り正確なものでないからで、それで、結果はわずかにずれている。しかし、それでも長さが2倍になると抵抗が2倍になることを示している。

**さらなる実験：**

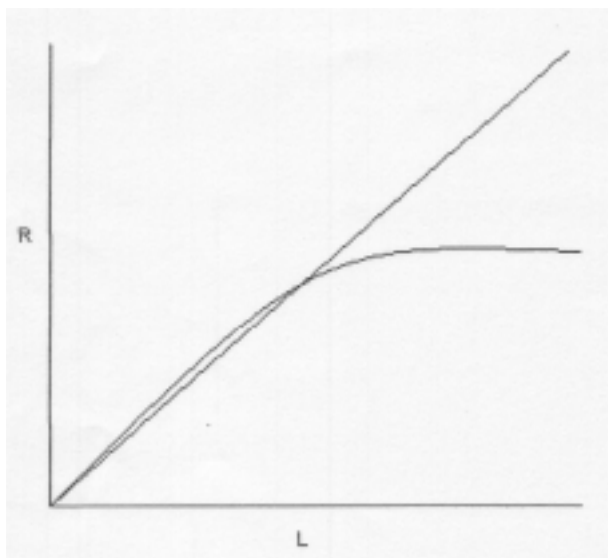
さらなる実験がさまざまなやり方で可能である。

ワニ口クリップの代わりに、異なる材質（炭素棒）を用いることができるだろう。これで実験がより正確になるだろう。

導線を一定の温度に保つことができる。そのために、導線をオイルバスにつければよい。もし、温度が一定であれば、導線の長さによって温度が変わって導線の抵抗が変わるのを心配しなくて済む。

また、より正確なグラフを得られるために、導線の長さを変化させることもできる。それによって、もっと多くの分析のための結果が得られるだろう。また、より長い長さを試してみることもできる。これは、長さを延長することで、線が直線が続けるのか、あるいは曲がって水平になるのかを明らかにしてくれるだろう。

E6b



A レベルコースワークの事例 (A レベル化学、P(「計画する」11 満点中 11 点)、I(「実施する」12 満点中 11 点、A(「証拠を分析し考察する」11 満点中 11 点)、E(「評価する」11 満点中 11 点)により、合計 45 点満点中 44 点)

(できるだけ生徒の記述と教師の書き込みに忠実になるよう訳した)

## 「メタノールからペンタノールまでの、一連のアルコールの燃焼熱の違いに関する研究」

### 目的

燃料としてのエネルギー効率を調べるという見地から、(分子の小さい順に) 5 番目までのアルコールについて燃焼熱を確定する。

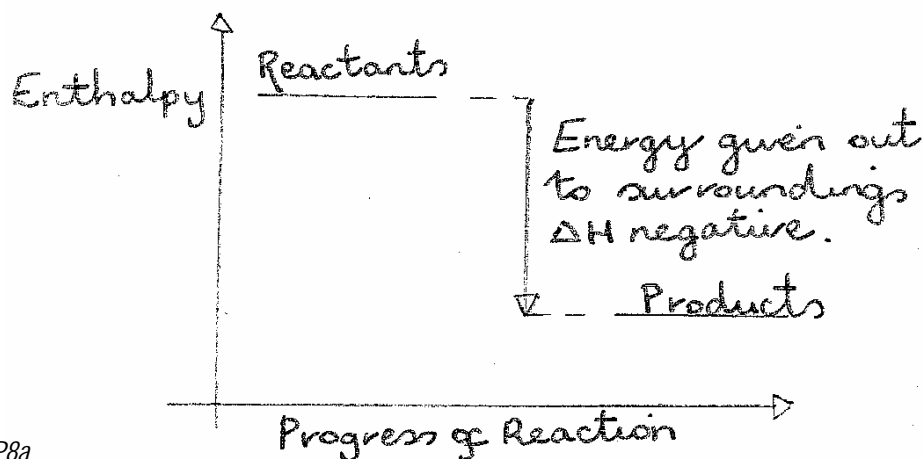
P5a

物質同士に反応が生じるときには、反応物と、その周りの環境との間に、熱交換が起こるが、それは発熱であったり、吸熱であったりする。この熱の交換が、エンタルピーの交換ということだ。そしてこの熱交換は、化学的な計算や実験によって求めることができる。ここで使う燃焼という意味は、燃料を酸素の中で完全に燃焼させるということである。

燃料は燃焼によって外部に熱を放出するが、また燃料自身をも加熱する。これは燃焼という反応が発熱反応であることを意味しており、すべてのアルコールの燃焼について当てはまる。

### 発熱反応

反応過程のはじめの段階は、化学結合を切断する段階で、その後新たな生成物を合成する段階になる。したがって、結合をきるときにはエネルギーを必要とし、生成物を合成するときには、エネルギーが作られる。反応の開始時においては、結合エネルギーをきるために、ある一定のエネルギーが必要になる。このエネルギーが、結合が形成されたときに生まれるエネルギーよりも低い場合には、余分なエネルギーが、外部に熱として放出される。この過程は、結合エネルギーのレベルを表した下の図のようになる。



P8a

余分な熱が失われ、エンタルピーエネルギーが減っていくような反応系は、発熱反応ということになり、エンタルピーエネルギーの値はマイナスになる。燃焼熱は下のように表すことができる：



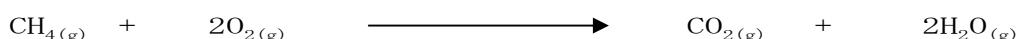
$\Delta H$  はエンタルピーの変化であり、'c' は「燃焼」を意味する。というのも、エンタルピーの変化はいろいろな反応によってもたらされるためである。たとえば化合物の生成による変化の場合には 'f' ということになる。 $\ominus$  という記号は常温常圧における反応ということを意味している。常温常圧とは、物質が安定な状態である、1 気圧、絶対温度 298 度ということだ。これによってエンタルピーの変化を比較可能なものとし、誤った結果に対して、一定の指標を与えることができる。

吸熱反応は、結合を切るエネルギーの方が、反応の最終段階で結合が生成されるときエネルギーよりも大きい反応である。したがって物質は外部からエネルギーを取り込まなくてはならないため、温度が下がるという結果になる。これはアルコールの燃焼には該当しない。燃焼は常に発熱反応である。

### 結合エネルギーの理論

以下にアルカン的一种であるメタン（類似のアルコールはメタノール）の燃焼について例示する。各結合エネルギーは（データ集の値によって）計算される。燃焼によるエンタルピーの総量は、結合の生成による値から、結合の切断による値を引くことにより求められる。

$$\Delta H_c = H_{\text{生成物質}} - H_{\text{反応物質}}$$



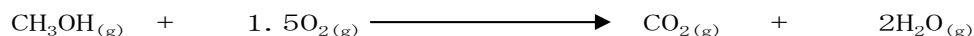
結合の切断			結合の生成		
		平均の結合エンタルピー			平均の結合エンタルピー
C-H	4 × 413	1652 kJmol <sup>-1</sup>	C=O	2 × 802	1610 kJmol <sup>-1</sup>
O=O	2 × 498	996 kJmol <sup>-1</sup>	O-H	4 × 464	1856 kJmol <sup>-1</sup>
	総量	2648 kJmol <sup>-1</sup>		総量	3466 kJmol <sup>-1</sup>

結合の生成におけるエンタルピーの総量のほうが大きいので、この反応は発熱反応となる。そしてこの二つの差がメタンの燃焼熱ということになる。-818 kJmol<sup>-1</sup> (3466 - 2648) 発熱反応であるから値はマイナスになる。反応において、多くのエネルギーが生み出されれば、この値は大きくなる。よって結合エネルギーの大きい物質が、燃焼において多くの熱を放出するということになる。アルコール類の中で効率のよい燃料を探るとこの実験で行おうとしている物質についても同様な計算を行うことができるが、詳細は後で述べる。

このエンタルピーの計算は、他の不確定な要因が影響を与えることなく、最良の状態での燃焼すると仮定しているが、実際にはいくら最小限にとどめようとしても、燃焼の程度に影響を与えるような数々の要因の影響があるだろう。

818 kJmol<sup>-1</sup>という数字は燃焼によるエンタルピーエネルギーから算出されたものである（周囲への熱の放出であるから負の値）。これはメタン 1mol がすべて完全に燃焼すれば、818 kJmol<sup>-1</sup>の熱が外部に放出されるということである。私が調べている 5 種類のアルコールも同じような仕組みで熱を放出する。アルコールが燃料とされるのはこのためである。

メタノール (CH<sub>3</sub>OH) の燃焼によるエンタルピーの変化はメタンで行ったように、以下のように計算される。結合エネルギーを以下のように計算した。



結合の切断		平均の結合エンタルピー	
C-H	3 × 413	1239	kJmol <sup>-1</sup>
O-H	464	464	kJmol <sup>-1</sup>
C-O	358	358	kJmol <sup>-1</sup>
O=O	1.5 × 498	747	kJmol <sup>-1</sup>
総量		2808	kJmol <sup>-1</sup>

結合の生成		平均の結合エンタルピー	
C=O	2 × 802	1610	kJmol <sup>-1</sup>
O-H	4 × 464	1856	kJmol <sup>-1</sup>
総量		3466	kJmol <sup>-1</sup>

$3466 - 2808 = -658 \text{ kJmol}^{-1}$  がメタノールの  $\Delta H_c^\ominus$  ということになる。

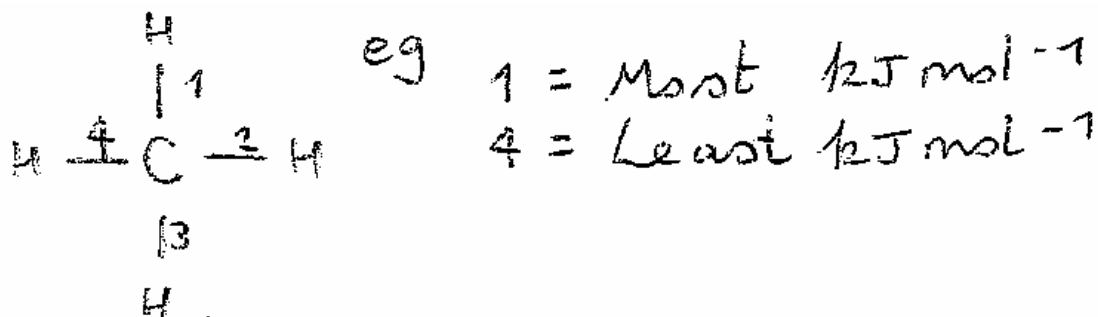
P11a

データ集の数字はこの計算した数字と少し違っている。それはこの計算で用いた水の結合エネルギーは気体の状態における値であるためだ。データ集では液体の状態での数値を使っている。

データ集によると、アルコールの燃焼熱の値は常温常圧下で、次のようになっている。

$\Delta H_c$ Methanol (メタノール)	CH <sub>3</sub> OH	-726.3	kJmol <sup>-1</sup>
$\Delta H_c$ Ethanol (エタノール)	C <sub>2</sub> H <sub>5</sub> OH	-1366.7	kJmol <sup>-1</sup>
$\Delta H_c$ Propanol (プロパノール)	C <sub>3</sub> H <sub>7</sub> OH	-2017.3	kJmol <sup>-1</sup>
$\Delta H_c$ Butanol (ブタノール)	C <sub>4</sub> H <sub>9</sub> OH	-2674.9	kJmol <sup>-1</sup>
$\Delta H_c$ Pentanol (ペンタノール)	C <sub>5</sub> H <sub>11</sub> OH	-3322.9	kJmol <sup>-1</sup>

以上のデータは各燃料の結合エネルギーの平均値を使って計算されたものである。というのたとえばメタンであれば、次の4つの結合はそれぞれ違ったエネルギーを持っているからだ。



上の表から、5種のアルコールの  $\Delta H_c$  の値が下に行くほど大きくなっていくことに気づくであろう。これについて以下に詳しく説明する。

Name of Alkane	Bond enthalpy of Products
<u>Methane</u> $\begin{array}{c} \text{H} \\   \\ \text{H}-\text{C}-\text{H} \\   \\ \text{H} \end{array}$ $\text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O}$	Products: Carbon dioxide and water $\text{C}=\text{O} \quad 2 \times 805 \quad 1610 \text{ kJ mol}^{-1}$ $\text{H}-\text{O} \quad 4 \times 464 \quad 1856 \text{ kJ mol}^{-1}$  $\text{Total} = \underline{3466} \text{ kJ mol}^{-1}$
<u>Ethane</u> $\begin{array}{c} \text{H} \quad \quad \text{H} \\   \quad \quad   \\ \text{H}-\text{C} - \text{C}-\text{H} \\   \quad \quad   \\ \text{H} \quad \quad \text{H} \end{array}$ $\text{C}_2\text{H}_6 + 3\frac{1}{2}\text{O}_2 \rightarrow 2\text{CO}_2 + 3\text{H}_2\text{O}$	$\text{C}=\text{O} \quad 4 \times 805 \quad 3220 \text{ kJ mol}^{-1}$ $\text{H}-\text{O} \quad 6 \times 464 \quad 2784 \text{ kJ mol}^{-1}$  $\text{Total} = \underline{6004} \text{ kJ mol}^{-1}$
<u>Propane</u> $\begin{array}{c} \text{H} \quad \text{H} \quad \text{H} \\   \quad   \quad   \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{H} \\   \quad   \quad   \\ \text{H} \quad \text{H} \quad \text{H} \end{array}$ $\text{C}_3\text{H}_8 + 5\text{O}_2 \rightarrow 3\text{CO}_2 + 4\text{H}_2\text{O}$	$\text{C}=\text{O} \quad 6 \times 805 \quad 4830 \text{ kJ mol}^{-1}$ $\text{H}-\text{O} \quad 8 \times 464 \quad 3712 \text{ kJ mol}^{-1}$  $\text{Total} = 8542 \text{ kJ mol}^{-1}$
<u>Butane</u> $\begin{array}{c} \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \\   \quad   \quad   \quad   \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{H} \\   \quad   \quad   \quad   \\ \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \end{array}$ $\text{C}_4\text{H}_{10} + 6\frac{1}{2}\text{O}_2 \rightarrow 4\text{CO}_2 + 5\text{H}_2\text{O}$	$\text{C}=\text{O} \quad 8 \times 805 \quad 6440 \text{ kJ mol}^{-1}$ $\text{H}-\text{O} \quad 10 \times 464 \quad 4640 \text{ kJ mol}^{-1}$  $\text{Total} = 11080 \text{ kJ mol}^{-1}$
<u>Pentane</u> $\begin{array}{c} \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \\   \quad   \quad   \quad   \quad   \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{C}-\text{H} \\   \quad   \quad   \quad   \quad   \\ \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \quad \text{H} \end{array}$ $\text{C}_5\text{H}_{12} + 8\text{O}_2 \rightarrow 5\text{CO}_2 + 6\text{H}_2\text{O}$	$\text{C}=\text{O} \quad 10 \times 805 \quad 8050 \text{ kJ mol}^{-1}$ $\text{H}-\text{O} \quad 12 \times 464 \quad 5568 \text{ kJ mol}^{-1}$  $\text{Total} = 13618 \text{ kJ mol}^{-1}$

P11a

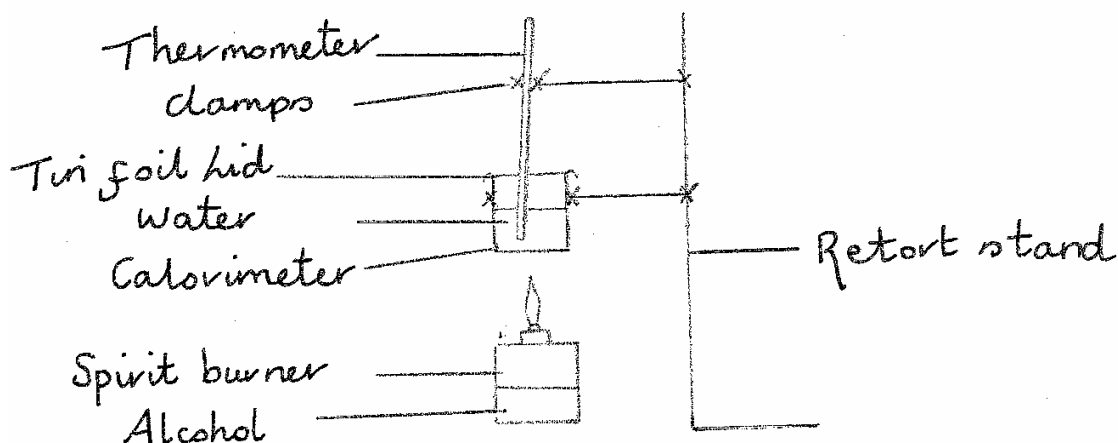
以上がアルカンにおける分子の構造と ( $\Delta H_c$  増加の) 傾向である。しかしこれを自分の研究であるアルコールの場合に適用することができる。生成物の結合エネルギー (結合の生成により、外部に出てくるエネルギー) が表の下になるにつれて大きくなっている。下のグループほど、炭素、水素、酸素原子が多く存在しているため、燃焼の際により多くの二酸化炭素や水が精製されるということであるから、下の方ほど能力が高い。ペンタノールの方へ行くにつれ、1モルあたり、より多くのエネルギーを放出するというのである。このことから、ペンタノールの方へ行くほどより多くの熱を放出し、メタノールがこの5種類のアルコールの中では最も能力が低い。

これは前出の表の中に、メタノールの燃焼熱が1モル当たり2596kJであり、ペンタノールに比べるとかなり低いということと一致する。

アルコールの燃焼について、理論は説明したが実験においてはこれらの値の精度を変化させる多くの要因がある。理論を実際に試してみるために、これらの要因について検討する。

## 方法

### 実験装置の図



アルコール類 メタノール、エタノール、プロパノール、ブタノール、ペンタノール

### その他の器具

- \* 液体を正確に測るメスシリンダー
- \* 0.01gまで測ることができる天秤
- \* 断熱のための陶板とアルミホイル
- \* ストップウォッチ

### アルコールの燃焼の精度に影響を与える要因

#### アルコールランプ

##### 芯の大きさ

P8b

太くて長い芯は表面積が大きくなり、炎に触れるアルコールの量が変わるので、芯の長さや直径はアルコールの燃焼に影響を与えるであろう。同じ太さで同じ長さの芯を使わなくてはならない。また、多種のアルコールが混合しないように、各アルコールにはそれぞれ別個の芯を用いなくてはならない。

P11b

##### ランプの大きさ

熱量計からの距離を等しく保つために、同じ大きさのランプを用いなくてはならない。そして一定の量を毎回測って入れなくてはならない。

##### 炎の高さ

銅の熱量計に炎があたると、一度に直接熱が伝わって急に水を温めるし、炎があたらなければ温まりにくいだろうから、この問題を軽減するために、常に一定の距離を保たなくてはならない。

### 温められる水

もし水の量が一定でないとすれば、たとえば水の量が多くなれば、 $20^{\circ}\text{C}$ 水温を上昇させるのにかかる時間が変わってくるので、多くのアルコールを消費してしまうことになる。

### 熱量計

P11b

材質が違くと、熱量が異なり、中の水の温度が同じになるまでの時間が変わってしまう。したがって同じタイプの熱量計を用いることが必要である。銅は熱量が小さく、すぐに温まって、水にエネルギーを与えるのにもっとも短い時間で済む。

### 熱の損失

この実験においてもっとも不確かなのは、すべての熱を比熱計に与えることができるかどうかということである。外部に熱が漏れないように十分な断熱を毎回施すことが必要である。

### 温度計

温度計の精度について不確かであれば、それが結果を変えてしまうことになるであろう。毎回同じ温度計を用いることが必要である。

### 予備実験

実際の実験の前に、不確定な要素をできるだけ減らし、どの測定方法が最も正確で適した結果を与えるかについて調べてみた。

### メタノール

P11a

燃焼前のアルコールランプの質量 =  $116.56\text{ g}$  燃焼後のアルコールランプの質量 =  $115.27\text{ g}$

水を  $20^{\circ}\text{C}$  上昇させるのに用いたアルコールの質量 =  $1.29\text{ g}$  ( $116.56 - 115.27$ )

水温の上昇 :  $19^{\circ}\text{C} - 39^{\circ}\text{C}$

かかった時間  $3.20$  分

水の体積 =  $65\text{ cm}^3$

芯の高さ =  $10\text{ mm}$

アルコールの量 =  $35\text{ cm}^3$

温度変化 =  $20^{\circ}\text{C}$

芯から熱量計までの距離 =  $8\text{ cm}$

芯の直径 =  $5\text{ mm}$

### メタノール 2

P11a

燃焼前のアルコールランプの質量 =  $116.00\text{ g}$  燃焼後のアルコールランプの質量 =  $114.94\text{ g}$

水を  $20^{\circ}\text{C}$  上昇させるのに用いたアルコールの質量 =  $1.06\text{ g}$

水温の上昇 :  $20^{\circ}\text{C} - 40^{\circ}\text{C}$

かかった時間  $2.04$  分

### 変更点

芯の高さ =  $15\text{ mm}$

芯から熱量計までの距離 =  $6\text{ cm}$

その他の条件は上の実験と同じにした。条件を適正なものにするために、同じ芯と、アルコールランプを用いた。上に示したような変更により、同じように熱するための時間が短くなり、少ないアルコールで済むようになった。このように改良した方法で、今後も実験を続ける。



メタノール 3

燃焼前のアルコールランプの質量=131.40g 燃焼後のアルコールランプの質量=129.57g

水を20℃上昇させるのに用いたアルコールの質量=1.47g

水温の上昇：20℃—40℃

かかった時間 1.23分

今回は断熱をよくするためにアルコールランプの周りに4枚の陶板を置き、アルミホイルでふたをした。これにより、外部に逃げる熱が減り、大部分の熱が熱量計に届くようになった。そして水の温度を20℃上昇させるのにかかる時間が短縮された。これは断熱によって、熱がよく伝わるようになったということを示している。

P11a

予備実験の結果より、つぎのような実験条件を定めて、実験を行うことにした。

水の体積=65cm<sup>3</sup>

温度変化=20℃

芯の高さ=15mm

芯から熱量計までの距離=6cm

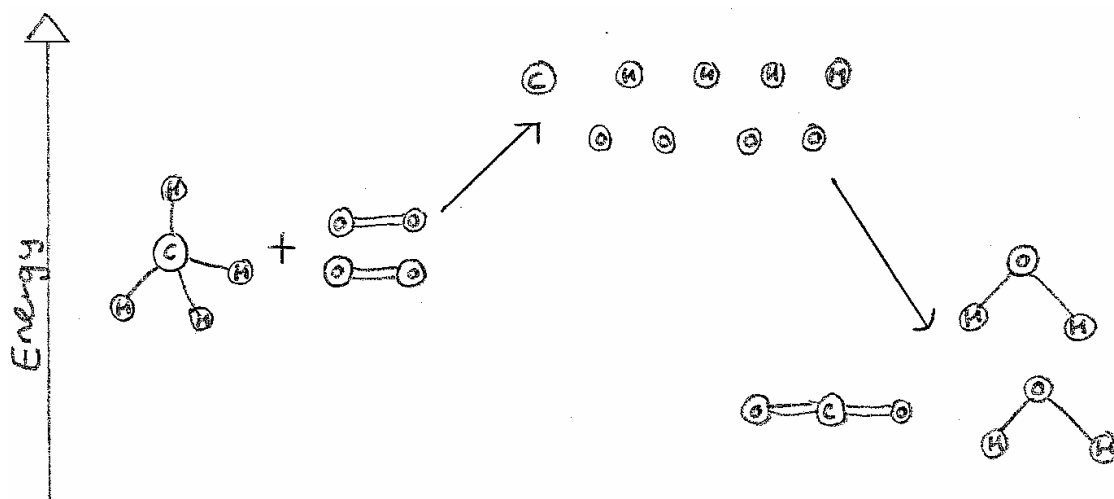
アルコールの量=35cm<sup>3</sup>

芯の直径=10mm

熱量計 銅製

断熱 陶板4枚 アルミホイル

第2次



上の図は化学の参考書にあるエンタルピーの部分からとったものであるが、メタンの燃焼にかかわるすべての分子が含まれている。反応は結合の切断と、結合の生成を含んでいる。メタンにおいては、炭素と水素の間の4つの単結合と、酸素原子の間の二重結合を切断する必要がある。発熱反応においては生成物が作られる際に、結合を切るのより大きなエネルギーが放出される。たくさんの原子が切断されれば、より大きなエネルギーが生まれる。よって1モルのペンタノールは、1モルのメタノールに比べてより多くのC=OやH<sub>2</sub>Oを作り出すので、炭素数の多いアルコールの方が、より多くのエネルギーを一定時間内に作り出すということになる。(以下実験の部分一部省略)

P11a

A2a  
A2b  
A5a  
A5b

上の結果は、水を20℃温めるのに、表の下のグループの方に行くほど、少ないアルコールで済むということを示している。それはまた、水を温めるのに、ペンタノールがもっとも短い時間で済み、メタノールが最も時間がかかるということも示している。これは私の予想通りで、5種のアルコールはデータ集にあったような傾向を示した。アルコールの（燃料としての）能力は、（炭素数の多い）ペンタノールへ向かうにつれ高くなる。これは炭素や他の原子を最も多く含み、多くの結合が生成することになるから、1モルあたり、より多くのエネルギーを放出するというので、少ない燃料で済む。計算によれば、ペンタノールが最大のエンタルピーを持ち、メタノールが最小、その他のアルコールはその中間に位置する。

次の方程式による。

$$E = m C \Delta T$$

ここでEは、一定量の水（60 cm<sup>3</sup>）の温度を20℃上昇させるのに必要なエネルギーであるから、常に一定となる。推測される（不確かな）要因の影響を最小限にした場合である。

$$\text{エネルギー} = \text{質量} \times \text{水の熱量} \times \text{温度変化}$$

kg                      J kg<sup>-1</sup> K<sup>-1</sup>                      °CまたはK

A8a

$$\text{水の密度} = 1 \text{ g cm}^{-3}$$

$$\begin{aligned} \text{したがって} \quad E &= 65 \text{ g} \times 4.2 \times 10^3 \times 20^\circ\text{C} \\ &= 0.065 \times 4.2 \times 10^3 \times 20 = 5460 \text{ J} = \underline{5.46 \text{ kJ}} \end{aligned}$$

したがって、水の温度を20℃上昇させるのに、5.46 kJのエネルギーが必要である。

それぞれのアルコールの燃焼熱を計算するために、水の温度を20℃上昇させるのに何モル必要であったかを計算しなくてはならないので、

エネルギーをモル数で割って、

$$\frac{\text{使用したアルコールの質量}}{\text{アルコールの分子量}} = \text{モル数} \quad \frac{\text{水の温度を20℃上昇させるのに必要なエネルギー}}{\text{使用したアルコールのモル数}} = \text{燃焼熱}$$

これらの式からそれぞれのアルコールの能力を算出することができる。

メタノールの能力

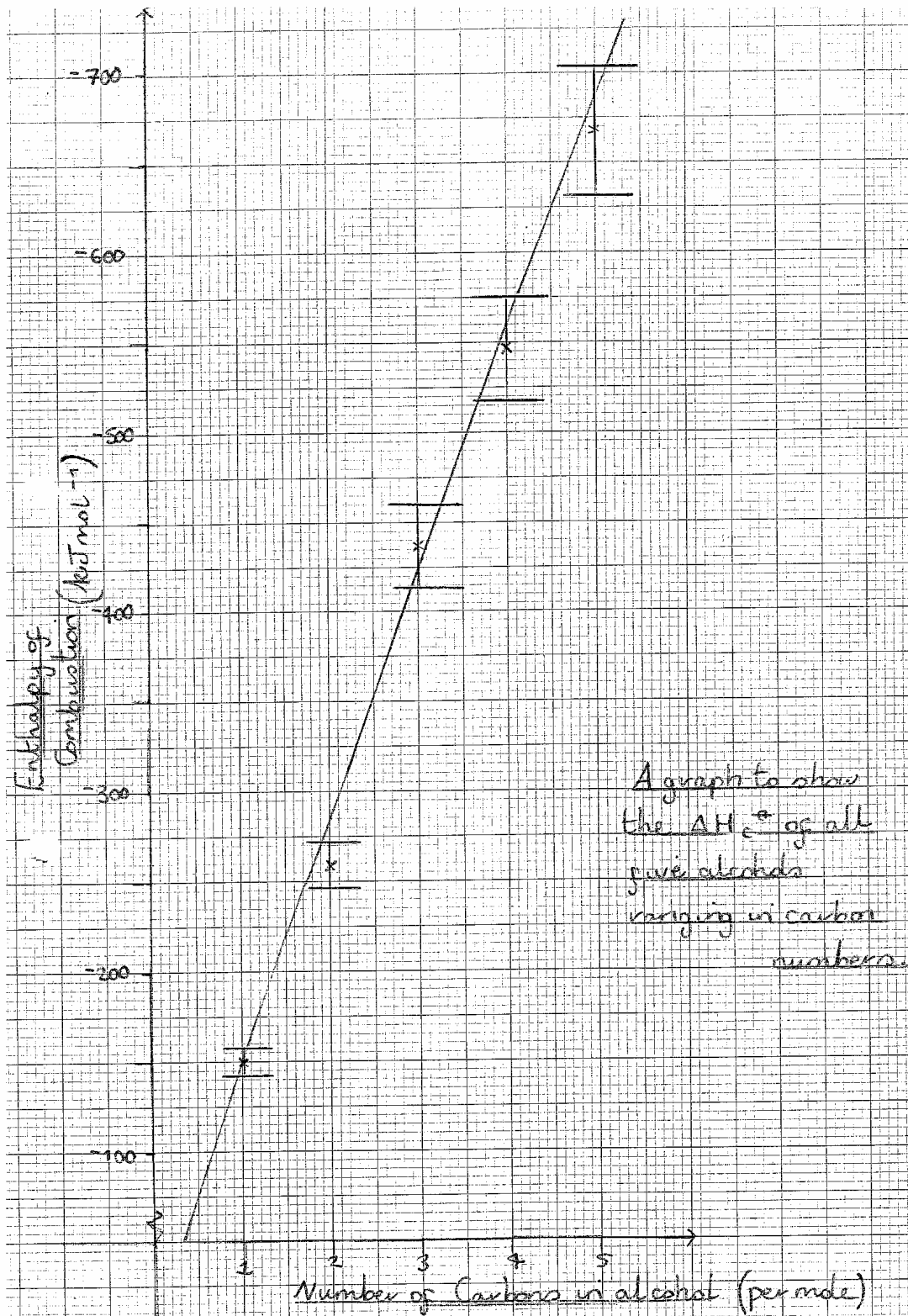
分子量：CH<sub>3</sub>OH

炭素 = 12

$$12 + (3 \times 1) + 16 + 1 = 32.0 \text{ g}$$

水素 = 1

酸素 = 16



$$\text{燃焼したメタノールのモル数} = \frac{1.16 \text{ g}}{32.0 \text{ g}} = 0.0363 \text{ モル} \quad (3 \text{ s f})$$

$$\text{メタノールの燃焼熱} = \frac{5.46 \text{ kJ}}{0.0363} = 150 \text{ kJ mol}^{-1} \quad (3 \text{ s f})$$

A11a

これと全く同じ計算過程を5種類全てのアルコールに適用して、これらのアルコールの燃焼熱を比較するために、表にまとめた。燃焼熱が高いほど、能力の高いアルコールということである。なぜなら、1モルあたり、より多くの熱を発生させるからだ。

アルコール	燃焼させた質量	燃焼させたモル数	燃焼熱
メタノール	1.16 g	0.0363	-150 kJ mol <sup>-1</sup>
エタノール	0.97 g	0.0211	-259 kJ mol <sup>-1</sup>
プロパノール	0.75 g	0.0125	-437 kJ mol <sup>-1</sup>
ブタノール	0.74 g	0.0100	-546 kJ mol <sup>-1</sup>
ペンタノール	0.72 g	0.00818	-667 kJ mol <sup>-1</sup>

予想されたように、ペンタノールに向かっていくにつれ、燃焼熱は増加する。数値は全て負の値であるが、これは熱量計の中の水に、熱を与えているためである。

- A11a 次ページのグラフは、アルコールの炭素数と、先ほどの実験結果から計算したアルコールの燃焼熱との関係を示したグラフである。各点を最もうまくつなぐ回帰直線は、燃焼熱が炭素数に比例して大きくなることを示す直線になっている。ペンタノールは最も大きい燃焼熱を示し、メタノールが最も低い。これは第2次における、アルコールの構造からペンタノールが一番高いであろうという予想を支持している。各点は直線の近くに分布しており、結果が正確であることを示している。これら5種のアルコールの燃焼熱はデータ集に載っているものとずいぶん違っている。これは燃焼に影響する要因、特に熱の損失という要因の影響を受けていないことに起因する。これは理想的な環境における理論値で、自分の実験に比べるとずいぶん高いものになっている。これは、教室という理想的ではない環境のために、かなり多くの熱が外部に漏れてしまったためである。しかしながら、その次のグラフに示すように、(理論値と)自分の実験を比較してみると、自分の実験は他の(不確かな)要因に影響を受けているものの、(炭素数と燃焼熱の)関係は同じ傾向を示していることがわかる。
- A11b
- A11a

### 誤差計算

実験結果に影響を与える要因があるという点からは、この実験が理想的なものではないと考えられるが、実験に含まれる誤差を計算してグラフ上に示すことで、本当に的確な回帰直線になっているか検討する。もしもこの直線が的確であれば、この結果は変則的なものではなく、値をばらつかせる何らかの要因に影響を受けたのだということがわかる。

各アルコールの $\Delta H_c$ を計算するときの誤差について計算する。

$$E = m C \Delta T$$

水の質量の誤差  $-1 \text{ ml} = 1 \text{ cm}^3 = 1 \text{ g}$

E2a 温度測定の誤差 =  $0.5^\circ\text{C}$

E2b これらの誤差は、実験者の不正確さによるものである。

E5a Cの値は、データ集に載っていたものなので、誤差を無視できる。

E11b

Eの計算に係る誤差の総計

$$E = (0.065 \text{ kg} \pm 0.001 \text{ g}) \times 4.2 \times 10^3 \times (20^\circ\text{C} \pm 0.5^\circ\text{C})$$

これをパーセントに直して、

$$E = (0.065 \text{ kg} \pm 1.54\%) \times 4.2 \times 10^3 \times (20^\circ\text{C} \pm 2.5\%) = 5.46 \text{ kJ} \pm 4.04\%$$

したがってエネルギーを求める際の誤差の割合は、 $5.46 \text{ kJ} \pm 4.04\%$

燃焼したアルコールのモル数を求める際の誤差

$$\text{モル数} = \frac{\text{使用したアルコールの質量 (g)}}{\text{アルコールの分子量}}$$

天秤からくる質量の誤差:  $0.01 \text{ g}$

アルコールの分子量は、データ集のものなので、誤差を無視できる。

各アルコールのモル数を求める際の誤差:

アルコール	燃焼させた質量	燃焼させたモル数
メタノール	$1.16 \text{ g} \pm 0.862\%$	$0.363 \pm 0.862\%$
エタノール	$0.97 \text{ g} \pm 1.031\%$	$0.0211 \pm 1.031\%$
プロパノール	$0.75 \text{ g} \pm 1.333\%$	$0.0125 \pm 1.333\%$
ブタノール	$0.74 \text{ g} \pm 1.351\%$	$0.0100 \pm 1.351\%$
ペンタノール	$0.72 \text{ g} \pm 1.389\%$	$0.00818 \pm 1.389\%$

全体の誤差割合を求めるために

$$\frac{5.46 \text{ kJ} \pm 4.04\%}{\text{修正されたモル数}}$$

例えば

$$\text{メタノール: } \frac{5.46 \text{ kJ} \pm 4.04\%}{0.363 \pm 0.862\%} = 150 \text{ kJ mol}^{-1} \pm 4.902\% = 150 \text{ kJ mol}^{-1} \pm 7.353$$

**A11b**

不確かさの絶対量が計算されている。

したがってメタノールの燃焼熱の範囲は、 $143 \text{ kJ mol}^{-1} - 157 \text{ kJ mol}^{-1}$

5種のアルコールについて全く同様な計算をして、燃焼熱の誤差の範囲を表にまとめた。

アルコール	$\Delta H_c$ の誤差範囲
メタノール	$-143 \text{ kJ mol}^{-1}$ $\longrightarrow$ $-157 \text{ kJ mol}^{-1}$
エタノール	$-246 \text{ kJ mol}^{-1}$ $\longrightarrow$ $-272 \text{ kJ mol}^{-1}$
プロパノール	$-414 \text{ kJ mol}^{-1}$ $\longrightarrow$ $-460 \text{ kJ mol}^{-1}$
ブタノール	$-517 \text{ kJ mol}^{-1}$ $\longrightarrow$ $-575 \text{ kJ mol}^{-1}$
ペンタノール	$-631 \text{ kJ mol}^{-1}$ $\longrightarrow$ $-703 \text{ kJ mol}^{-1}$

これらの誤差範囲は先ほどのグラフ上に赤い線でその範囲を示した。

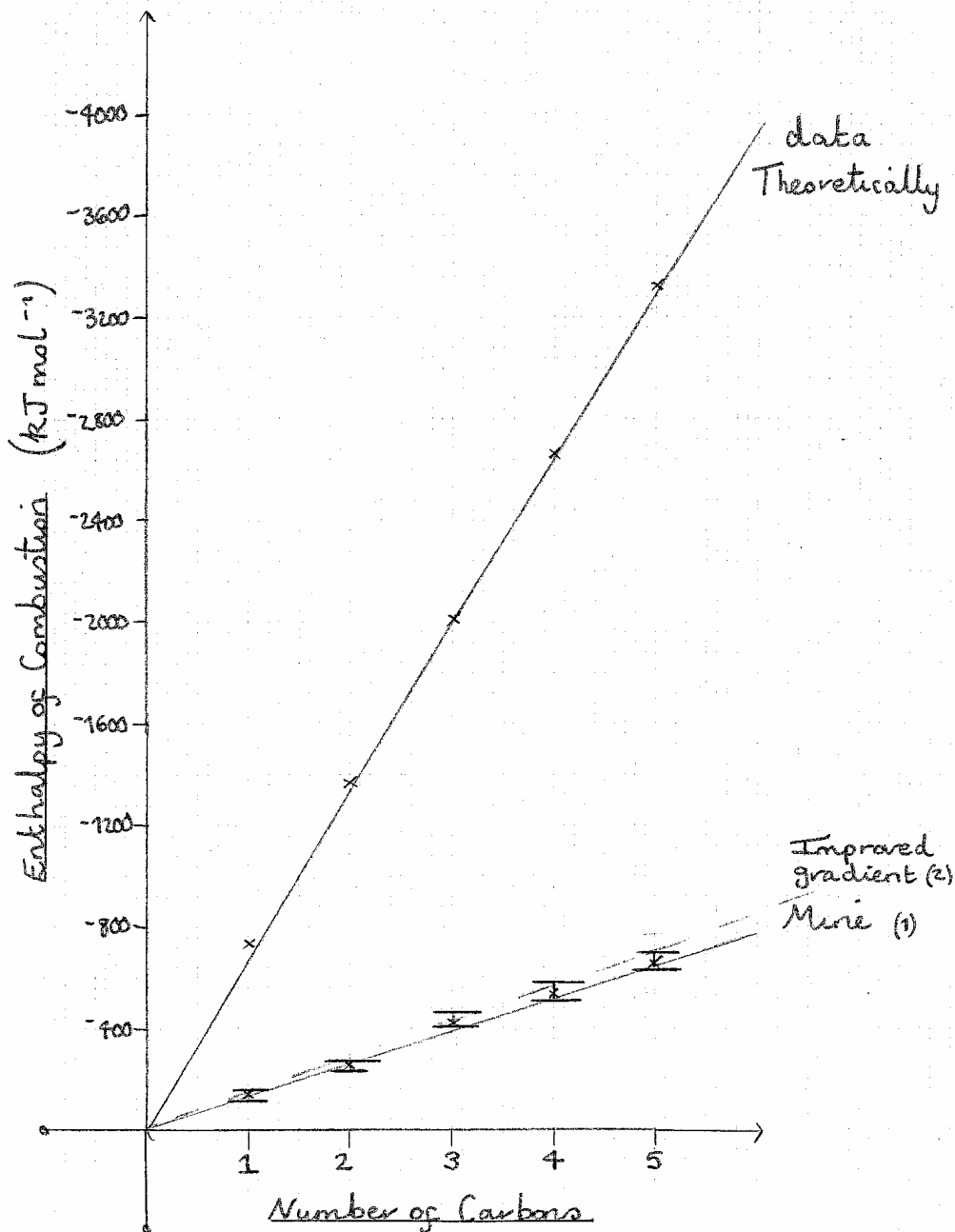
**E2b** グラフに示されるように先ほどの回帰直線はエタノールを除いてこのエラーバーの間を通過している。これはほんの少しのずれであるとはいえ変則的な結果であり、エタノールの実験が何か不確かな要因の影響をより強く受けたということになろう。

**E11b** グラフの傾向は、十分に比例の関係を示しており、炭素数が増えるにしたがって燃焼熱が増加し、アルコールの能力が高まっているといえる。ひとつの結果がわずかにずれているが、回帰直線は予想を支持している。実験は成功しており、燃焼率の影響を受けているとはいえ、ほぼ矛盾のないものになっている。自分の実験結果をデータ集の理想的な値と比較してみる。ここでは、燃焼は外部の影響を受けずに理想的な状態で燃焼させている。私の結果は、その値よりも低いものになっているが、ペンタノールに向かうにつれて、燃焼熱が増加するという傾向は一致すべきものである。これは私の実験が、全ての熱が熱量計の容器に伝わっていないということの影響を受けたものだ。主な原因は、外部への熱損失である。実験の評価のところで、これを改善する手立てについて詳しく述べる。データ集の燃焼熱と私の実験のデータの差から、熱が水に伝わる効率が悪かったことがわかる。これらの本当の値は、熱をうまく閉じ込めて測定されたものだが、これについても後で詳しく述べる。

**E5b**  
**E8b** 次のグラフは、自分の実験結果と、理論値との比較をしたものである。

もし自分の実験の値が非常に正確であったら、値は低いにせよ、これらの直線は平行になるはずである。そしてこれら2本の直線は同様な傾向を示すことになるであろうし、同じ形態になることだろう。これは全ての（エタノールからペンタノールまでの）実験に対して、常に一定量、燃焼熱を引き下げる要因が働いたときのみ起こりうる。しかしこの実験ではそのようになっておらず、人的な不正確さか外部環境要因によって、その要因（が与える影響の大きさ）は実験の途中で変化している。二つのグラフはどちらも直線になっており、燃焼熱が比例的に大きくなっていくことを示している。しかし自分の実験では水や熱量計に対してそれほど熱を与えていない。このグラフにエラーバーを付加すると、新たな回帰直線を引くことができる。これは私の実験結果の誤差範囲を考慮に入れたものだが、そのようにするとグラフの傾きは急になり、本当の値に近づく。

A Graph to Show the comparison  
of my experiment with  
the data book values



A17b とはいえ、依然としてこの二つのグラフはかなり異なっている。しかしこれは私の実験の熱損失によるものである。自分の実験は炭素数が1つ増えるごとに、燃焼熱が平均  $150 \text{ kJ mol}^{-1}$  ずつ増加していることがわかる。これが一定になっているために、グラフは直線になっている。燃焼熱の値は高くないけれども、燃焼熱の増加はある一定の傾向を持っており一貫している。

この傾向は本当の値についても見られ、炭素数が増えるごとに、燃焼熱が平均  $600 \text{ kJ mol}^{-1}$  ずつ増加している。自分の実験は、(不確かな要因の) 影響を受けない標準的で正確な燃焼熱と同じ傾向を示しているということだ。

### 評価

グラフで示したように、自分の実験は十分に一貫した結果を示している。しかし、データ集にあるほどの正確さには達していない。これは実験を行った環境が理想的なものではなく、外部からの影響を受けたためだと考えられる。データ集にある5種類のアルコールの燃焼熱は、アルコールが最も効率よく燃焼し、最大の熱を発生させるような理想的な環境下で測定されたものである。この(不確かな要因による) 影響は、燃焼熱の値が低くなり、炭素数の増加に伴う燃焼率増加の割合が低くなるという形でグラフに現れている。これは実験途中のさまざまな要因の影響によるものである。

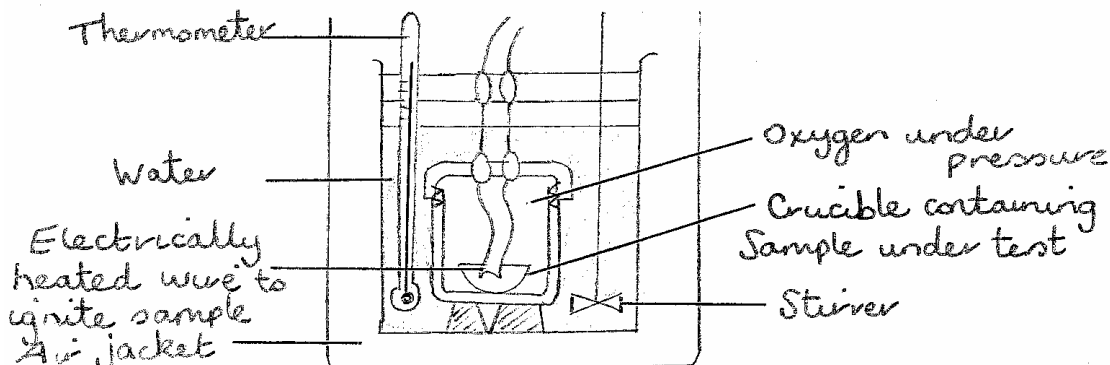
影響の大きさの順に述べる。

#### 1) 熱損失

E17b

アルコールランプの炎によって作られた熱は、いくつもの理由で、水まで十分に到達しない。このことは自分の実験結果が低い数値になったことでわかる。陶板を使って断熱することが熱損失を防ぐのに有効なことは予備実験の段階で明らかになった。しかし依然として、炎は水温を高める前に、空気や銅製の熱量計を温めなくてはならない。最小の時間で容器が温まる、最も比熱の低い材料でできた熱量計を用いたのはそのためである。また、すきま風や対流は熱損失を促進してしまうであろう。これらを取り除くことができれば(水の) 温度をあげることに使われる熱の量は大きく違ってくるであろう。考えられる方法としては、「爆弾熱量計」を用いることである。これは教科書にあるアイデアだが、より正確に熱を捕らえることができる。これは压力容器の中で、電熱線により着火するものである。すきま風もないし、空気の対流も起こらない。また、近接しているものへ熱が逃げることもないので、熱の損失がない。この方法は、一定の圧力、一定の体積のもとで行わなくてはならない。

E11a





## 2) ランプの芯

E11a

同じアルコールを燃焼させるのには同一の芯を用いたが、アルコールの種類を変えるときには、アルコールが混ざり合わないよう芯を交換しなくてはならなかった。新しい芯は、違った表面積を持っているだろうし、表面積が広ければより多くのアルコールが（炎に）接することになるから、変則的な実験結果をもたらすであろう。いくつかのアルコールの実験では、芯から出る炭素のカスが残ったが、それが次の実験に影響を与えたとも考えられる。したがって、毎回新しい芯に取り替えることにより、この要因は除外できる。芯の高さも、この実験結果の一貫性に影響を与えうる。というのも、芯が高くなれば炎は熱量計に近づくので、それによって多くの熱を与えることになったり、銅の容器の下につくすすの形成要因になったりする。このことは、熱量を変化させ、熱と容器の間で邪魔をすることになる。そして、水を温める前にすすを温めることになるので、時間が長くなる。アルコールの種類によってすすの生成量が異なっていた。これは不完全燃焼を示すものであり、周囲の酸素濃度を高めることで解消できる。自分の実験では最後の2つのアルコールが最も多くのすすを生成した。このことが、水温を20℃上昇させるのに長い時間がかかったことの原因かもしれない。

## 3) 反復実験

各アルコールは不確かさを取り除くために3回ずつ繰り返して測定をした。しかしさらに繰り返すことで、より正確な平均を導くことができるだろう。また、ヘキサノールを加えるといった、より多くのアルコールを実験してみることで、真の値とのより広範な比較が、可能になるであろう。

## 4) 予備実験

実験の理想的な状況を探るために、より多種の実験装置の考案や、新しいアイデアの導入が考えられる。

## 5) 測定

実験の誤差率は、測定の不正確さを許容しているが、メスシリンダーを容積の小さいものに替えたり、シリンジやピペットに替えることが精度を上げるのに有効であろう。水の体積の違いは、その熱量に影響を与え、一定の水温上昇にかかる時間を変化させることになる。温度計の位置も温度を決定するのに影響を与えるであろう。例えば、温度計が熱量計に触れていたならば、それは水の温度ではなく、銅の温度を測ってしまったことになり、大変異なったものになるだろう。したがって、温度計は容器の真ん中で、容器に触れないようにしておかなくてはならない。ストップウォッチを止めて、測ったときの温度が、正確な値ではなかったかもしれない。これが結果に影響していることも考えられる。これはより正確な温度計を用いることによって解消できる。

## 6) アルミホイルのふた

予備実験において、水温を上昇させるのにかかる時間が減ったことから、アルミホイルのふたが断熱に効果があることがわかった。しかしこれが水や空気の対流を妨げ、水全体の温度が様ではなかったとすれば、温度計の読みが不正確になったということも考えられる。

## 7) 質量の測定方法

自分の実験では、水を温めるのにどれだけのアルコールが使われたか、その質量を測定し、エネルギーの算出を行っている。正確な測定のためには、何千グラムも測れるような上皿天秤があると良い。

アルコールは異なった構造を持ち、下のグループに行くほど、多くの炭素を含んでいる。

この物理的な特性が、アルコールの燃え方に影響を与えているのかもしれない。アルカンにおいてその沸点は、炭素鎖が長くなるにつれて高くなる。炭素数が増えると、分子の中の分子間力が強まるためである。このことがアルコールにも当てはまるのではないだろうか。ペンタノールの方へ行くにしたがって、今ある結合を切ることが難しくなる。そのことが自分の実験における、ブタノールとプロパノールの燃焼熱が低くなってしまったことの説明にならないだろうか。これら二つのアルコールが燃焼するのに気化しようとする、他のアルコールより多くのエネルギーが必要になる。したがってより長い時間がかかり、アルコールの使用量は増えてしまう。この特徴は（分子量の大きい）アルコールにとって不利に働くが、より正確な実験で回避できる。

この実験は、データ集にもあるような、燃焼熱の傾向を明らかにした。かなりの割合の熱損失によって、燃焼熱の値は低くなってしまったが、（5種の）アルコールは比例的な燃焼熱の増加を示した。結果は、精度や、それが示す傾向から、一貫したものであり、不確かさは非常に低い。（回帰直線の）勾配は、本来の傾きより緩やかではあるものの、比例的に変化している。燃焼熱の傾向と、それが分子構造とどのように関連しているかが示され、実験結果によって証明された。いくつかの要因は他の要因に比べて、より大きな影響を燃焼に与えているが、前に示したように、要因は全実験を通じて一定なので、規則正しいグラフとなっている。

教師によるコメント：評価シートより

**P** 数多くの場面で、**11a/11b** をすべて満たしている。優れたレベルの理論で、上手く述べられている。正確性を要することを考慮して全体を計画している。  
(11点満点中、11点)

**I** ノートを参照  
(12点満点中、11点)

**A** 理論が散漫な感じもするが、実験結果とその傾向に関する深い考察やその前の理論との比較がそれを相殺している。  
(11点満点中、11点)

**E** 誤差がきちんと計算されている。誤差の各要因について分析され、改善方法を示している。  
(11点満点中、11点)

合計 44点 (最大 45点)

## 第 3 章

Science as a context for the Science of Thinking

「思考に関する科学」の文脈としての「科学」

フィリップ・アデイ教授（ロンドン大学キングス・カレッジ校）



## **Science as a context for the Science of Thinking**

**Philip Adey, Centre for the Advancement of Thinking, King's College London**

### **1 Introduction**

This paper will describe the origins, structure, and effects of an innovation called Cognitive Acceleration through Science Education (CASE). As the title indicates, CASE is based on the science *of* thinking, that is psychological theories of how we think and how our thinking can be improved, and it uses science in the school curriculum as the vehicle to enhance thinking.

CASE is designed as an intervention in the science curriculum of students aged about 11 to 14 years. It is not a complete curriculum, but its activities replace 'regular' science activities about once every two weeks. It had its origins in work done in the 1970s at Chelsea College in London which showed that many of the concepts included in science curricula in the United Kingdom (and throughout the world) made demands beyond the intellectual capability of the students for whom it was intended.

The team at Chelsea College, led by Professor Michael Shayer, took a scientific approach to the problem of difficulty. On the one hand, we needed an accurate description of the intellectual profile of the school population and on the other, we needed a way of measuring and describing the level of difficulty of science concepts. The theory of cognitive development which had been elaborated by Piaget provided us with just the sort of description we needed. Drawing on his descriptions of types of thinking available at different stages, we (1) developed an instrument with which curriculum materials could be analysed for the cognitive demands that they made, and (2) developed group tests of cognitive development (Shayer, Wylam, Küchemann & Adey, 1978) and used them in a very large scale survey to establish the levels of thinking of children at different ages in the school population of England and Wales. It was shown clearly (Shayer & Adey, 1981) that the science curriculum made unrealistic demands on the ability of many – perhaps a majority – of the school population.

There are in principle two possible approaches to this problem: make the science curriculum easier, or raise the intellectual capability of the students. While the former would be relatively easy, it would inevitably engender academic and political difficulties, and in any case may be seen as a defeatist solution. Although the prospect of raising all students' ability to think may appear daunting, this was precisely the aim of the CASE project, initiated in 1982.

## 2 The Underlying Psychology

By 'Cognitive Acceleration' we mean the process of accelerating students' 'natural' development process through different stages of thinking ability, towards the type of abstract, logical, and multi-variate thinking which Piaget describes as 'formal operations'. Formal operational thinking is characterised by the ability to hold a number of variables in mind at once - for example to be able to weigh up two sides of an argument, to consider even-handedly the advantages and disadvantages of a particular course of action, or to be able to see both the separate and combined effects of a number of input variables (for example, sunlight, carbon dioxide, water) on an outcome (the production of glucose). Piaget had suggested that this type of thinking becomes available to children as a process of natural intellectual development around the ages of 14 or 15 years. However our Chelsea survey showed that only 30% of 16 year olds were capable of such thinking, and this conclusion was supported by work with college freshmen in the US, and parallel (but smaller scale) surveys in other parts of the world.

Up to the late 1970s, attempts at cognitive acceleration had shown little evidence of success. However, these studies had adopted a rather short-term and direct instructional approach, as if the mind's ability to process information could be changed by learning a new set of rules. We believed that such approaches were flawed since the mind's processing ability grows slowly, in response to demand placed upon it by challenging problems. This gives us the first of five 'pillars' of CASE theory, *Cognitive Conflict*. This occurs when a student encounters a problem which he cannot easily solve for himself but which, with carefully structured help from an adult or more able peer can either be solved, or will lead to gain in understanding of the nature of the problem. The principle of cognitive conflict is also encapsulated within the idea of a 'Zone of Proximal Development' (ZPD) developed by Vygotsky (1978). The ZPD is the difference between what a child can do unaided, and what he can do with the help of an adult. Vygotsky says '... the only good learning is that which is in advance of development'. In other words, learning tasks which are well within the child's capability do not provide the challenge which stimulates cognitive growth. CASE activities are designed to be intellectually challenging.

Vygotsky also provides us with the second 'pillar' of CASE: *social construction*. As a good Soviet psychologist, Vygotsky was very conscious of the fact that people work together to build knowledge; it is a social process. We talk to one another, argue, listen, and gradually build our understanding through social interaction. CASE teachers learn to facilitate good social interaction, creating a classroom atmosphere in which polite disagreement is encouraged, and it is all right to change one's mind.

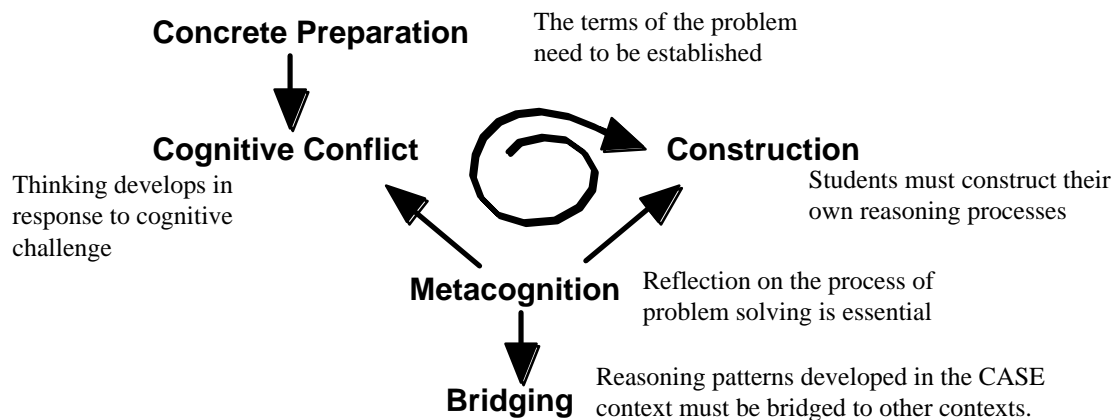
The third pillar of CASE theory is the encouragement of *Metacognition*. Metacognition means essentially 'thinking about your own thinking', although as an extremely fashionable

notion in cognitive psychology the word metacognition has been used in many different ways (Brown, 1987). We can only help ourselves to develop higher level thinking if we take some control of our thinking, that is, become conscious of ourselves as thinkers. In CASE, students are encouraged to take time to reflect on how they solved a problem, what they found difficult about it, what sort of reasoning they used, how they sought help and what sort of help they needed. This is time consuming and quite difficult to do, and teachers and students need a lot of help and encouragement initially to become more metacognitive in their approach.

There are just two more pillars of CASE theory. One is the idea of *Concrete Preparation*. You cannot simply present students with a difficult problem and expect the cognitive conflict to do the work of cognitive acceleration. There must be a phase of preparation in which the language of the problem is introduced, along with any apparatus to be used and a context in which the problem is set. The aim is to ensure that the difficulties encountered are just intellectual, and as far as possible are not confounded by problems language or context. The final pillar is *Bridging*, the linking of ways of thinking developed in the particular context of the CASE activity to other contexts within the science, mathematics, or other parts of the curriculum and to experiences in real life. If it is to become generally available, reasoning developed within a special context must be abstracted, and the student shown how it can be used as a general thinking tool.

Figure 1 illustrates the relationship of these five pillars to one another.

**Figure 1: Five ‘Pillars of CASE Wisdom’**



The relationship of cognitive conflict to social construction, shown by a spiral arrow, is not straightforward. When faced with a problem, we tend to seek simple solutions. We ‘short-circuit’ a full analysis of the problem in order to reach an accommodation which will meet the immediate

needs of the situation. For example, in determining what factors cause iron to rust and finding that nails in water rust faster than dry nails, the student will be content with the solution 'rust is caused by water' without looking more deeply into the possible effects of air as well. Cognitive conflict by itself does not automatically lead to re-construction of concepts or to reaching a full understanding. The cognitive conflict must be maintained and this can only be done by the teacher through close questioning. This gives a hint about the nature of the pedagogy required for cognitive acceleration which will be described in a later section on professional development.

The 'Five Pillars' provide a foundation for the pedagogy of cognitive acceleration, but by themselves they specify nothing about the subject matter context. Teaching methods based on the Piaget-Vygotsky foundation outlined above could be developed in any subject matter. So why did we choose to work through science rather than, say, mathematics, history, or English? There was a pragmatic element to the answer - the early Chelsea work which led to the CASE project was science based and both Michael Shayer and myself had science backgrounds. But there was also a good theoretical reason for at least starting the work in science. The original detailed description of formal operations provided by Inhelder & Piaget (1958) is characterised by a set of mental 'schemata': control of variables, ratio and proportionality, compensation, equilibrium, correlation, probability, and the use of formal models. These schemata are immediately recognisable by scientists and science teachers as descriptive of important types of relationships between variables, and they are the stuff of experimental design and the elucidation of general patterns of behaviour in the natural world. Formal operations are a quite general way of processing data in any intellectual field and the schema of formal operations can be interpreted in the context of any academic subject area, but their application to science is fairly straightforward. Science presented itself as a most obvious gateway into the development of high level thinking.

When we started developing activities in 1984 we had not yet fully articulated the theoretical model outlined above. The schemata of formal operations were established as the framework about which activities would be structured, the 'pillar' of cognitive conflict was recognised as central to the process of cognitive acceleration and constructivism had always been a main pillar of Piaget's account of cognitive development. The need for concrete preparation was a pragmatic necessity which came from our experience as teachers, and bridging likewise seemed of obvious importance if the schema were to be generalised. But our elaboration of the importance of metacognition grew throughout the project, from being implicit in the type of questioning that we promoted, to becoming an explicit and very important part of the CASE method. This gradual evolution of the 'pillars' of CASE as a complete theoretical structure underpinning the design and delivery of activities has since become very important in the process of professional development of teachers, to be described later.



### **3 Development of the Curriculum Activities**

#### *Target population*

In line with the origin of CASE described in section 1, we were concerned with a broad range of ability, the majority of the student population for whom science appeared to be rather difficult. In terms of ability, our target was the middle eighty to ninety percent of students. Now, the experience of cognitive conflict will depend on an individual's ability. What provides an interesting and productive puzzle for one individual may appear trivial to a more able child and incomprehensible to a less able peer. While careful design of activities and flexible pedagogy can provide a wide range of levels of conflict within a particular activity, we considered it impracticable to include within our target population either the exceptionally able child, who would already be using formal operations by the age of 11 years, or those with serious learning difficulties who at 11 years may still be preoperational.

We targeted the 11 - 14 year age range because for the great majority of students this is the age of preparation for formal operational thinking. There is some evidence (Epstein, 1990) that there are brain-growth spurts at about 11 in girls and 12 in boys which may be part of a physiological maturation programme evolved to prepare adolescents for the intellectual demands of adulthood. Our survey of the population referred to in the introduction showed that only a small proportion of actual children attained the ages of cognitive development described by Piaget in his 'epistemic subject'. The population survey may be read as an indication of a deficit in the quality of stimulation provided for the majority of children at home and in school. On this reading, such a deficit should be remediable by appropriately designed stimulation at the right ages.

There is a pragmatic reason also for choosing 11 - 14 years as the age of operation of CASE. In the UK, 11+ is the age of transfer from primary school, which have class teachers who teach all subjects, to secondary school with specialist subject teachers. An intervention set within a science context would require science teachers who already understood - implicitly if not explicitly - the nature of the scientific reasoning patterns which form the context of the intervention.

Armed with the main features of a theoretical model, with the schemata of formal operations and with our experience as science teachers, the CASE authors (initially Michael Shayer, then joined by myself and Carolyn Yates) started to draft activities we thought would be appropriate for our target population. We discussed the form and practicalities of the activities amongst ourselves and with other academics and teachers. We ourselves taught each of the drafted activities to classes in London comprehensive schools which represented the age and ability

range of our target population, and the kind of social and ethnic mix typical of inner city schools in the UK. After a year of the funded project we had a bank of some twenty activities ready for a wider trial. Here two of the activities will be described to illustrate the application of the ‘pillars’ to practice. The materials are published as ‘Thinking Science’ (3<sup>rd</sup> edition: Adey, Shayer, and Yates, 2001)

**TS4: Tubes.** This is the fourth activity in the programme. In the previous activities, the ideas of variable, values of variables, and relationships have been introduced. Students have a box of small tubes. Questioning in a whole class discussion ensures that they identify the variables and values: length of tube (short, medium, long); width of tube (wide or narrow); and the material of the tube (copper or plastic). This is the *concrete preparation* phase of the activity, familiarising students with the basic ideas they are going to manipulate and the practical apparatus they are to use. Now they are asked to tap the tubes, open end against the palm of their hands, and listen to the note produced. The question is this: what affects the note that you get? They have some free exploration time and are asked, if they think they know what affects the note, to explain to the teacher or to another student what they think and why they think it. There is often a need, after some minutes, to call the class together and suggest that they take tubes just two at a time.

This is the phase of *cognitive conflict* and *social construction*. A child may come up with the claim that the width of tube affects the note. ‘Show me’ says the teacher. The student demonstrates with two tubes of different width that produce different notes. Looking at the tubes, teacher points out that they also have different lengths. ‘How do you know whether it is the length or the width that affects the note?’ Here the teacher is establishing some cognitive conflict, challenging the student to take account of a variable which she had not yet noticed. Typically a child might answer ‘both width and length that affect the note’. She does this as it seems a simple way to resolve the conflict, but the teacher perseveres with the questioning, concluding ‘go and choose another pair of tubes, but this time try to find a pair that will give us a clear answer’. Note that the teacher does not direct the student to choose two tubes in which only one variable has altered. The whole point is that the student must construct for herself this control of variables strategy.

In a mixed ability class of 12 year olds, it is possible that there will be one or two children who find the whole task so easy that they do not experience much cognitive conflict. For these the teacher may suggest a higher level task, such as looking for interaction between variables. There may be one or two others who, at the end of the 60 or 70 minute lesson, remain quite confused by the whole exercise and still fail to see the point of controlling variables. The great majority, however, will have experienced (through interaction with the apparatus, worksheet questions, the teacher and with other students) sufficient conflict to have constructed for

themselves at least the beginning of a control of variables strategy. The full development of this into an internalised, unconscious, schema which is 'naturally' brought to bear on all experimental situations will still take some time, but essential groundwork has been laid and the concrete 'change everything and see what happens' schema will have been severely shaken, if not broken up altogether. Even for the least able students who remain confused at the end of the activity there will have been a struggling with the problem and some doubts cast on the ineffective concrete strategy. Even a slight sense of unease at the way in which experimental questions are approached is of value. It is the cognitive struggle which is critical in the promotion of cognitive development, so the objective has been reached if every child experiences some cognitive conflict and goes some way towards finding a resolution satisfactory to her or himself. Toward the end of the lesson, the teacher asks students to reflect on mistakes they have made, on what they have learned, and gets them to re-trace their reasoning as they struggled with the control of variables strategy – this is the *metacognitive* work. Finally, they are invited to think of other topics in which the general control of variables strategy may be useful, to *bridge* the thinking from this lesson to other contexts.

**TS 18, Treatments and Effects**, is taught in the second year of the programme. This is set in the context of the schema of correlation. The *concrete preparation* discussion concerns two researchers who are testing the effect of a new fertiliser on the growth of carrots. Each has a treatment and non-treatment set of carrot plants and counts the number of plants in each set which show increased yield over a standard. The data is presented to the whole class as two 2 x 2 tables (treated / untreated and shows effect / no effect) and the discussion emphasises that the data from the untreated carrots is as important as that from the treated carrots (concrete operators tend to look only at treated carrots, to see if many show increased yield). Each group of students is now given a set of 20 cards. Each set of cards shows one organism (rose, wheat, cow, pig, or sheep), and also shows whether or not the organism has received some treatment (e.g. fertiliser, pills to make more milk, etc.) and whether the animal or plant demonstrates an effect (by growing more, producing more milk, meat, etc.). Students first sort the cards into four piles according to whether they have:

- A not been treated and not shown an effect
- B not been treated but shown the effect anyway
- C been treated but not shown an effect
- D been treated and shown an effect.

Students then address the question of whether any effect seen is likely to be the result of the treatment or not. For example, IF the treatment cause the effect, in which of the four piles A, B, C, and D would you expect to find large numbers? This generates considerable *cognitive conflict* and discussion in groups (*social construction*) leads to the conclusion that you would expect piles A and D to be large, and B and C to be small. In the discussion of these results the terms positive correlation, negative correlation, and no correlation are introduced to help students think about what sorts of relationships exist between treatments and effects. This activity models at a simple level the type of experimental evaluation of treatments which is at the heart of much medical, agricultural, and other research. Without an understanding of correlation and associated probabilistic relationships, the majority of popular science reports in newspapers are incomprehensible. Again, a *metacognitive* discussion explores the difficulties that students encountered, and they are encouraged the *bridge* the idea of correlation to topics outside the science classroom..

In none of these ‘Thinking Science’ lessons do students complete notes of ‘conclusions reached’ or ‘knowledge captured’. There may be no written product at all, as the worksheets are used just to record data which forms the raw material for thinking about relationships. This again highlights a difference between the CASE intervention activities and the regular science curriculum, a difference which some teachers initially find hard to accept.

### *Fitting it into the curriculum*

CASE does not offer a complete alternative science curriculum. Indeed, the pedagogic difficulty of managing intervention lessons and the fact that no science content is explicitly covered make it unsuitable as a substitute for regular science teaching. Furthermore, while the uncertainty with which students are sometimes left at the end of CASE lessons is productive in moderate amounts, it might well become demotivating if it were a permanent feature of science lessons. CASE is described as an ‘intervention’ both because it is a process of intervention in ‘normal’ cognitive development, but also because it is an intervention in the regular science curriculum. CASE activities are taught instead of regular science activities once every two weeks which might represent about twenty percent of the time allocated to science. Teachers sometimes say ‘it sounds like a good idea but we do not have the time for it’. This is an understandable position, but the reality is that very little time is actually ‘lost’ to the curriculum content material. This is partly because CASE already covers some of the process objectives of the curriculum, but mostly because as the students’ thinking develops so they are able to understand and make sense of the regular curriculum material more efficiently, in less time. Luckily we have very good evidence to

support this claim, and this evidence generally persuades teachers that the risk of 'losing' so much curriculum time is worth taking, at least on a trial basis.

#### **4 Trials and Evaluation**

The effect of the CASE intervention on students' cognitive development and academic achievement determined from our original research project has now been widely reported - see for example Adey & Shayer (1993, 1994); Shayer & Adey (1992a, 1992b) . A summary of that work will be given here before considering more recent evidence.

##### **THE 1984-87 EXPERIMENT.**

The results which will be described here are for the ten experimental classes in seven schools that continued with the programme for a period of two years. In each of these schools, one or two classes were designated as 'experimental', and from September 1985 started to use the Thinking Science activities once every two weeks for two years. Four of the experimental groups had children aged about 11+ years, and six had children aged about 12+ years. In each school also parallel 'control' classes were identified which were matched with the experimental classes for age and ability. The control classes were taught their regular science curriculum without loss of time for the CASE intervention.

All classes were given a pre-test of cognitive development and then at the end of the two year intervention period, post-tests of cognitive development and a test of science achievement. This was the end of the intervention programme, but one year later we revisited the schools to collect information on all of the students' science achievement. One further year later, in July 1989, those classes which had started the CASE intervention in their Year 8 took their General Certificate of Secondary Education (GCSE) examinations. This is the national public examination taken at 16 years by all students in schools in England and Wales. For all of the students who had previously been in classes designated as experimental and control we collected the grades attained in science, mathematics, and English. One year on again (July 1990), those who had started in Year 7 sat their GCSEs and again we collected their grades. We thus had the data which allowed us to compare (a) cognitive growth and (b) academic achievement over a long period of initially matched students some of whom had experienced the CASE intervention and some of whom had simply followed their regular science courses.

**Table 1: Residualised gain scores on successive tests after completion of two year CASE intervention, based on pre-cognitive tests September 1984**

		Group	Number	Mean gain	Standard deviation	Significance, p<	Effect size (s.d.)
Immediate post cognitive test		11+ boys	29	-0.21	0.95	-	-
		11+ girls	27	0.08	1.10	-	-
July 1987		12+ boys	65	0.70	1.00	.001	0.75
		12+ girls	52	0.03	0.98	-	-
1 year delayed science achievement		11+ boys	37	2.72	15.45	-	-
		11+ girls	31	7.02	12.76	.025	0.60
July 1988		12+ boys	41	10.46	16.6	.005	0.72
		12+ girls	36	4.18	14.41	-	-
GCSE 1989	Science	12+ boys	48	1.03	1.34	.005	0.96
		12+ girls	45	0.19	1.38	-	-
	Maths	12+ boys	56	0.55	1.23	.005	0.50
		12+ girls	54	0.14	1.27	-	-
	English	12+ boys	56	0.38	1.27	.05	0.32
		12+ girls	57	0.41	0.96	.01	0.44
GCSE 1990	Science	11+ boys	35	-0.23	1.46	-	-
		11+ girls	29	0.67	1.36	.025	0.67
	Maths	11+ boys	33	-0.21	1.59	-	-
		11+ girls	29	0.94	1.26	.005	0.72
	English	11+ boys	36	0.26	1.65	-	-
		11+ girls	27	0.74	1.32	.025	0.69

All results will be reported in terms of residualised gain (r.g.) scores, determined from the regression of any post test on to the initial pre-test. Note that r.g. scores build in comparison to controls and that by definition the mean r.g. score of a control group must be zero. Table 1 summarises results for boys and girls at different ages of starting. There are a number of striking features of these results :

- The immediate effects seem to be rather limited, but (1) more recent immediate effects obtained on cognitive development have been much larger (see below) and (2) there is a strong correlation on an individual student basis between cognitive gains over the two year intervention programme and subsequent gains in GCSE scores.
- There is a long term, and apparently growing, effect of the intervention of students' academic achievement. The effect of the raised cognitive levels will be, starting at the end of the intervention, to improve student's ability to benefit from normal classroom instruction. Such improvement is likely to be cumulative as better understood conceptual learning provides a sounder platform for further learning, and so on.
- There is a strong 'far transfer' effect. An intervention programme set in a scientific context effects students' achievement in mathematics and in English literature. Such transfer implies that CASE has tapped into and influenced a deep-seated function of the mind which has a broad effect on students' intellect.
- There seems to be an age/gender interaction effect, in that the intervention is most effective with younger girls and with older boys. Although this notion fits neatly with a model of a cognitive window of opportunity for the promotion of formal operations, which in line with their generally earlier maturity at this age comes earlier for girls than it does for boys, we must be very careful before drawing such a conclusion. For one thing, the 11+ group was actually more able overall than the 12+ group, both age groups starting the intervention at about the same mean level of cognitive development. For another, more recent data does not show anything like the same gender effect.

## MORE RECENT RESULTS

In the original research experiment we were able to measure effects on experimental classes against well-matched controls. However, the numbers were relatively small, we ourselves were still in the process of inventing the method for training the teachers, and the teachers themselves were working on the project in isolation within their schools.

Following the publication of the long-term effects on GCSE scores in May 1991, there was a great demand from schools for the materials and methods that would enable them to replicate the results. Since then, we have been running a series of two-year inservice teacher education courses to introduce the methods. This professional training will be described in more detail in Section 5. Although we are now collecting much new data, an important difference between this and the original experiment is that now we have a method which we believe works, we cannot ethically deny it to any class just to provide an experimental control. One way of analysing new data is to compare gains made by CASE schools with the national norms established in the Chelsea survey

(Shayer and Adey 1981, and see Section 2). From the first cohort of schools participating in the CASE training programme, we were able to collect pre- and post-test data on levels of cognitive development for sixty-three classes in eight schools. Some of these classes made a Year 7 (aged 11+ years) start on the intervention, some a Year 8 (12+) start, and one school started the intervention in both years. A summary of the effect sizes of the school mean residualised gain scores compared with national norms is shown in table 2.

**Table 2: Effect sizes of cognitive development: residualised gain scores in eight schools which participated in CASE training, 1991-93**

School	Start age	Effect size ( $\sigma$ units)
1	11+	0.67
1	12+	0.76
2	11+	0.69
3	11+	1.12
4	11+	1.12
5	12+	0.80
6	11+	1.0
7	12+	0.29
8	12+	1.26*

\* by comparison with previous year 9 group, questionable.

Of the sixty-three classes, there was a significant negative effect in one class, possibly due to some error in the administration of the pre-test. In four others there were insignificant negative effects. In three classes there were positive effects of less than  $0.3\sigma$ . In all of the remaining fifty-five classes there were significant positive effects of the CASE intervention on children's rate of cognitive development. As we have shown previously, cognitive gains attained over the intervention period are related to subsequent academic gains.

In 1995 and 1996 we were able to collect data on academic achievement of CASE schools, compared with non-CASE schools, for the 'Key Stage 3 National Curriculum Test' (KS3 NCT), as well as for GCSE grades of students who had used CASE in 1991 – 93 and those who did CASE in 1994-96. I will present the KS3 results first and then the 1999 GCSE data. In the UK,



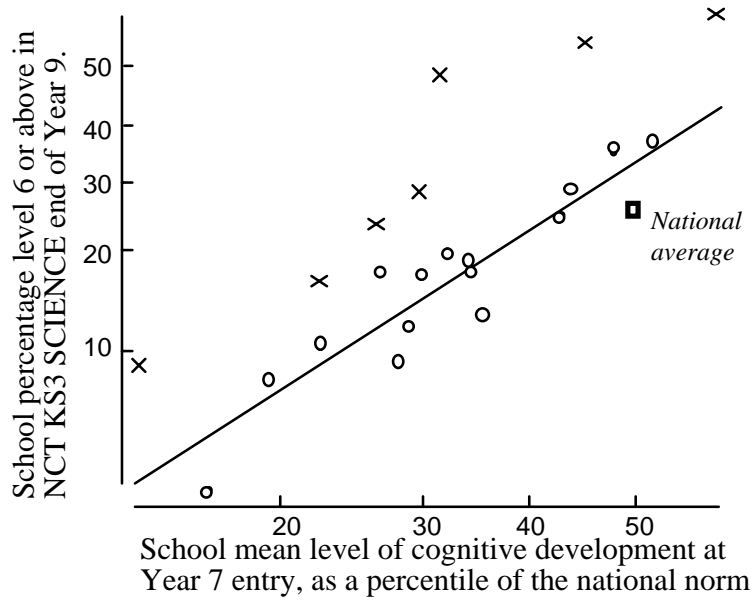
the government has instituted a series of nationally moderated tests to be given in various subject areas at the end of each 'Key Stage' of education, which means at the ends of years 2, 6, and 9 when children are about 7, 11, and 14 years old respectively. For schools which use Thinking Science in years 7 and 8, the KS3 NCT given at the end of year 9 provides a convenient measure of academic achievement one year after the end of the intervention.

In figures 2a, 2b, 2c each point represents one school. The horizontal, x, axis is the mean score of the school's students at the beginning of Year 7 (secondary school entry) on measures of levels of cognitive development, expressed as a percentile of the national average. This is a measure of the school's intake ability, which is a reflection of factors such as the socio-economic conditions in the school's environs and whether there are selective schools in the area which cream off the more able students. It so happens that almost all of the schools for which we have data at present are in the lower half of the intake ability range. The vertical, y, axis is a measure of success in the KS3 NCT. These tests are scored for National Curriculum levels, which fall on a range from 1 to 10 (or more recently 1 to '8 and over'). The percentage of students attaining level 6 and above at Key Stage 3 is commonly taken as a measure of the success of the school. In order to make the plot linear, all scores have been transformed into logits:  $\ln(\%/100-\%)$ . This is why the axis scales are not equal-interval.

These figures show, not surprisingly, that success on the Key Stage 3 tests is strongly related to the intake ability of the school's pupils. What is striking is that CASE schools all lie above - often far above - the regression line for control schools. This means that whatever the intake level of the school, CASE schools are adding significantly more academic value to their students than non-CASE schools. The effect is equivalent to an addition of about 30 percentile points to school mean academic achievements.

**Figure 2: relationship between school entry cognitive levels and NCT KS3 test performance at end of year 9 for CASE (X) and non-CASE (O) schools.**

**2a: Science**



**2b: Mathematics**

