

# 異なる研究アプローチが従属変数の選択へ与える影響

## Differences in selecting dependent variables between hypothesis-driven and data-driven research approaches

松室 美紀<sup>†</sup>, 三輪 和久<sup>‡</sup>

Miki Matsumuro, Kazuhisa Miwa

<sup>†</sup> 立命館大学, <sup>‡</sup> 名古屋大学

Ritsumeikan University, Nagoya University

matumuro@rm.is.ritsumei.ac.jp

### 概要

本研究は、データ分析を行う際の、従属変数の選択に焦点を当てている。特に、特定の仮説に基づき分析を行う仮説駆動アプローチと、分析結果から仮説を構築するデータ駆動アプローチの影響を検討した。実験の結果、仮説駆動アプローチを用いると、データを加工し、仮説の検証に適したデータを自ら作成するが、類似した種類のデータを分析に用いることが示された。一方、データ駆動アプローチを用いることは、様々なデータの分析につながるが、データの加工は促進しない。

キーワード: 科学的研究プロセス (scientific research process), データ分析 (data analysis), 仮説駆動 (hypothesis-driven), データ駆動 (data-driven)

### 1. はじめに

科学的研究は、一般に、実験の計画、実施、実験結果の分析、分析の結果に基づく考察からなる。本研究では実験結果の分析に焦点を当て、異なる研究アプローチがどのような分析の違いを引き起こすか検討を行った。

#### 1.1 仮説と独立変数の選択

多くの先行研究では、特定の仮説に基づき実験を計画、実行する、仮説駆動アプローチが、科学的発見や学習に有効であるとされている [1, 2, 3, 4]。それらの研究で用いられた課題は、参加者が実験を繰り返し、規則性や事実を発見することを目標とする課題であった。仮説駆動アプローチを用いることにより、それら一連の実験で、仮説と対応する変数を体系立て操作することが可能である。そのため、先行研究では仮説駆動アプローチの利点が最大限に生かされ、効率的な発見、学習へつなげたと考えられる [1, 2, 4, 5]。

しかし、特に社会学の分野における実験や調査では、そのコストを考えると、連続的な複数回の実施は困難である。そのため、1回の実験、調査において、様々な変数の値が測定され、それらの分析に基づき結論が導出される。分析の内容により、結論の内容や説得力が変化するため、いずれの変数をどのような方法を用い分析するかが非常に重要である [6]。

仮説駆動アプローチの利点は、主に実験の計画にあることを踏まえると、このようなデータ分析においては異なる効果を持つ可能性がある。そこで、本研究では、近年注目を集めているデータ駆動アプローチを用いた場合と、データ分析中の行動の比較を行う。データ駆動アプローチは、特定の仮説を持たずに、データの分析を行い、その結果に基づき、理論や仮説を構築するアプローチである [7, 8]。遺伝学等の様々な分野で、データ駆動アプローチによる発見がなされている。

#### 1.2 仮説空間とデータ空間

科学的発見の過程は Klahr らにより仮説空間と実験空間の二重空間の探索として論じられている [3, 9]。しかし、この枠組みには、データ取得後の分析の過程は含まれていない。そこで、本研究では、データの分析の過程を、仮説空間とデータ空間の二重空間の探索過程として考える。図1に概念図を示す。

##### 1.2.1 データ空間

データ空間は、分析に使用可能な全てのデータを含む。ここで、データとは従属変数の候補とその値の組合せと定義される [7, 10]。

データ空間の探索は統計的検定等を用い、各データを分析することにより行われる。従属変数としてその値を分析に利用された変数のみが探索されたと定義さ

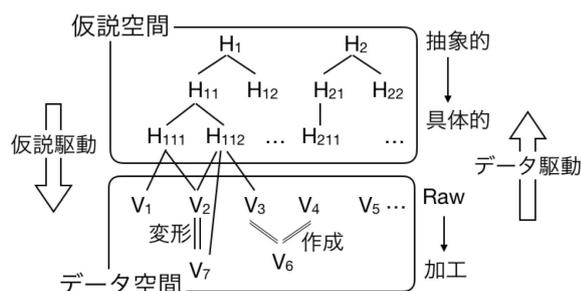


図1 仮説空間とデータ空間. H は仮説を, V は変数を表す.

れ, 利用されなかった変数は, 探索されなかった範囲とされる.

もし, 研究者が特定の変数や類似した変数ばかりを対象としてデータ分析を進めている場合, データ空間中の似通った部分のみの探索を行っていることとなる. そのため, 探索が偏っている, つまり, 探索の幅が狭いと定義される.

さらに, 研究者は変数の値を加工したり, 複数の変数の値を組み合わせるにより, 直接的に値を測定できない変数を検定に用いることが可能である. 例えば, 変数の値に対数変換を施したり, プレ, ポストテスト得点の差異から得点変化量を作成したりすることにより, ローデータを分析する以上の情報が得られる.

このように, ローデータを加工することにより, 多角的な検証を行ったり, 特定の主張に最も適した変数を用いた検定を行うことができる. このような加工したデータを用いたデータ分析は, データ空間の深い部分の探索に当たり, ローデータを用いたデータ分析はデータ空間の表層的な探索に当たる.

### 1.2.2 仮説空間

仮説空間は全ての仮説を含み, それらの仮説は抽象性に基づき階層構造をなしている [2]. 抽象的な仮説は一般的で多くの事象を説明するが, そこから特定の変数の値の傾向を予測することはできない. 一方, 具体的な仮説は, 厳密で少数の事象しか説明できないが, 特定の変数の値の傾向を予測する.

## 1.3 研究アプローチと2つの空間の探索

我々は, データ空間探索の幅と深さに関する仮説を導出, 検証した.

### 1.3.1 仮説駆動アプローチ

仮説駆動アプローチを用いた場合, 仮説空間からデータ空間の方向へ探索が行われる. 研究者が, 特定の仮説をターゲットとすると (e.g., 図1における  $H_{11}$ ), その子に当たる仮説が検討の対象とされる (e.g., 図1における  $H_{111}, H_{112}$ ). その際のデータ分析には, それらの仮説の検証に必要な変数のみ (e.g., 図1における  $V_1, V_2, V_7$  等) が従属変数として分析に利用される. つまり, データ空間の探索の幅は狭くなる.

また, 特定の仮説を持つため, その仮説を支持する可能性のある変数を予測し, 分析に利用しようとする. 予測した変数がローデータに存在しない場合は, データの加工を行い新しい変数を作成し, データ空間の深い探索を行うと考えられる.

### 1.3.2 データ駆動アプローチ

データ駆動アプローチを用いた場合, データ空間から仮説空間の方向へ探索が行われる. 研究者は, 興味や目的に基づきデータ空間全体から分析対象とする変数を選択する. そのため, データ空間の探索は制限されることなく, 広い探索が実施される. その結果に基づき, 具体的な仮説や理論が選択され, それらに基づきより抽象的な仮説や理論が形成される.

研究者は, まずは手元に揃っているデータから興味深い結果を見つけようとするため, データの加工の必要性は低い. つまり, データ空間の探索は浅くなると考えられる.

### 1.3.3 本研究の仮説

上記をまとめると以下の2つの仮説となる.

**深さ仮説** 仮説駆動アプローチはデータ駆動アプローチよりも深い探索を促進する.

**幅仮説** データ駆動アプローチは仮説駆動アプローチよりも広い探索を促進する.

合わせて, 仮説駆動アプローチをとった際に, 仮説がデータにより支持されるか否かにより, データ空間の探索が変化するかを検討する.

## 1.4 データ分析課題

参加者は, 仮説駆動アプローチを取る仮説駆動条件と, データ駆動アプローチを取るデータ駆動条件に分

表 1 データセット中の各変数.

カテゴリ	変数名	変数の説明
学習過程	使用時間	システムを使用した合計時間
	問題数	システムを使用して解いた問題数
	使用回数	システムを使用した回数
	不正解数	システム使用中に間違った回答をした回数
学習結果	全体得点	ポストテスト全体の合計点
	基礎問題得点	ポストテストの基礎問題の合計点
	応用問題得点	ポストテストの応用問題の合計点
質問紙	熟考度	質問「よく考えて問題を解いた」への回答
	動機付け	質問「意欲を持って自宅学習へ取り組んだ」への回答
	難易度	質問「システムを使うのは難しかった」への回答
基礎学力テスト	数学	数学の基礎学力テストの得点
	国語	国語の基礎学力テストの得点
学生調査	兄弟姉妹数	兄弟姉妹合計数
	携帯電話使用時間	1日に携帯電話を使用する時間
	塾の時間	一週間に塾で学習をしている時間

けられ、架空の実験で取得されたデータを分析するよう求められた。

#### 1.4.1 シナリオ

参加者は、2種類の数学の自宅学習システムを評価する実験で取得されたデータを分析することを求められた。参加者の課題は2つのサブタスクからなった。サブタスク1は、2つの内どちらの学習システムが学習効果が高いかを明らかにすること、サブタスク2は、なぜそのような学習効果の差が生じたかを検討することであった。

仮説駆動条件の参加者には、参加者自身が開発したとされる新システムの方が、従来の旧システムよりも学習効果が高いという仮説が与えられた。データ駆動条件の参加者には、システムの名称はシステム1、システム2とされ、優劣の予測は与えられなかった。参加者には、分析結果の考察に利用できるように、各システムの使用法が教示された。

学習効果の評価実験には、学力が同程度の30名の学生が参加したとした。各システムを用いた1週間の自宅学習中の行動データ、確認テストの得点とシステムに関する質問紙への回答が取得されたと教示を行った。

#### 1.4.2 データセット

参加者には表1に示した変数の値が与えられた。確認テストの得点は、仮説駆動条件の参加者の持つ仮説を支持するよう、全体得点と応用問題得点において、新システムの方が旧システムより有意に高くなるように設定された。

新システムの学習効果が高くなる理由として、「新システムを使用するとよく考えて問題を解いていた」という理由が導出されやすいよう、次の変数で新システムにおいて、平均値が有意に大きくなるよう値が設定された。熟考度と不正解数、そして、使用時間を問題数で割った一問あたりの思考時間である。ただし、使用時間、問題数のそれぞれの値では、2種類のシステム間に有意な差異が生じないように調整された。

#### 1.4.3 分析支援ツール

参加者には、分析対象とする変数を選択するだけで、統計的検定を実行できる分析支援ツールが与えられた。平均値の差異の検定(t検定)と相関の有意性の検定が可能であった。さらに、本ツールは2つの変数と演算子を選択すると、新しい変数を作成する機能を持った。

各条件の分析過程は、分析支援ツールにより統制された。仮説駆動条件の参加者には、はじめに仮説を記述させた後、仮説に基づき検定方法を記述、実施させ、最後に仮説が支持されたかを判断させた。

データ駆動条件の参加者には、検定方法の記述、実施の後に、なぜそのような結果が得られたかに関する仮説や解釈を記述させた。両条件ともに、参加者は一連の流れを繰り返し行った。

## 2. 実験 1, 2

2つの実験を合わせて記述する。変数の値は、実験1では仮説駆動条件の仮説を支持し、実験2では支持しないよう設定された。

### 2.1 方法

#### 2.1.1 参加者

実験1では、各条件23名ずつ計46名、実験2では、仮説駆動条件に23名、データ駆動条件に24名の計47名の学部生が参加した。

#### 2.1.2 手続き

はじめに、統計的検定とツールの簡単な説明を実施した。続いて、自宅学習システムと評価実験、データの教示を行い、30分間のデータ分析をサブタスク1から行わせた。参加者は、自由なタイミングで学習効果の検討の結果を入力し、サブタスク2を開始した。最後に、分析の結果からわかったことを自由記述させた。

実験1, 2の手続きは同一であり、差異は与えられたデータのみであった。先述の通り、実験1ではデータが仮説駆動条件で与えられた仮説を支持する。実験2では、実験1で新システムに割り当てられた値を旧システムへ、旧システムに割り当てられた値を新システムへと割り当てた。つまり、仮説駆動条件で与えられた仮説は支持されない。

## 2.2 結果

与えられたデータの内容と自由記述の内容が一貫している参加者のデータのみが分析に用いられた。実験1の仮説駆動条件から2名、実験2の仮説駆動条件から5名、データ駆動条件から1名の参加者が除外された。

#### 2.2.1 統計的検定の実施回数

図2に、全体、サブタスク1, 2における統計的検定の平均実施回数を示す。全体の実施回数は実験1ではデータ駆動条件の方が高い傾向が見られたが

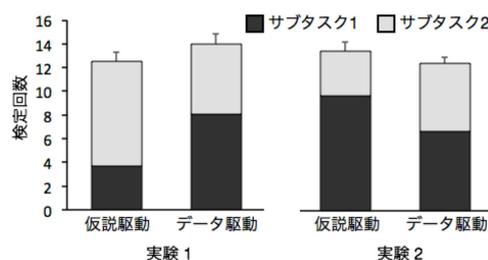


図2 統計的検定の平均実施回数 (エラーバーは標準誤差を示す)。

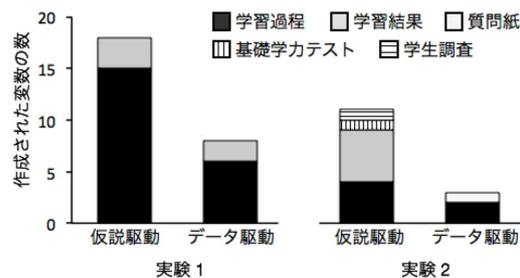


図3 作成された変数の種類。

( $t(42) = 1.304, p = .100$ ), 実験2では条件間に有意な差異はなかった ( $t(39) = 1.091, p = .141$ ).

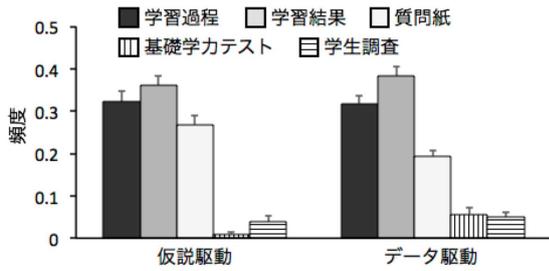
全体に占めるサブタスク1のための検定実施回数の割合を条件間で比較した。その結果、サブタスク1のための検定実施割合は、実験1では仮説駆動条件 (.280) よりもデータ駆動条件 (.584) において有意に多いが ( $t(42) = 3.858, p < .001$ ), 実験2では、逆にデータ駆動条件 (.538) よりも仮説駆動条件 (.717) において有意に多かった ( $t(39) = 2.228, p = .016$ ).

この結果は、自身の持つ学習効果に関する仮説が支持された場合は、学習効果の検討を素早く打ち切り、支持されなかった場合は長く検討を続けたことを示す。

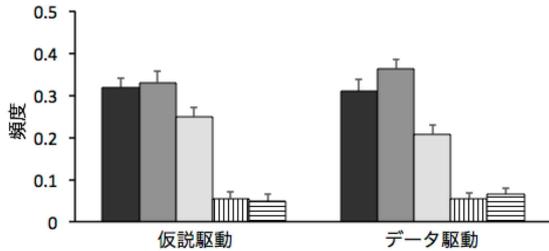
#### 2.2.2 探索の深さ

探索の深さの指標として、新しく作成された変数の数を用いた。両実験共に、仮説駆動条件において、データ駆動条件よりも有意に多くの変数が作成されていた (実験1, 0.851 vs. 0.348,  $t(42) = 2.343, p = .024$ ; 実験2, 0.600 vs. 0.136,  $t(39) = 2.310, p = .030$ )。ここから、仮説が支持されるか否かにかかわらず、仮説駆動条件の参加者の方がデータ空間の深い探索を行っていたことが示される。

図3は新しく作成された変数のカテゴリ (表1参照) を示す。分類は、作成に用いられた2つの変数のカテゴリ、および、参加者が命名した変数名に基づきなされた。仮説駆動条件に着目すると、学習効果に関する



(a) 実験 1 における検定頻度



(b) 実験 2 における検定頻度

図 4 各カテゴリに含まれる変数が従属変数として選択された頻度 (エラーバーは標準誤差を示す)。

仮説が支持された実験 1 では、多くの変数は学習過程へ分類された。一方で、仮説が支持されない実験 2 では、学習結果へ分類される変数が増加している。

### 2.2.3 探索の幅

表 1 に示される各カテゴリの変数が検定に用いられた頻度を算出した。その結果を図 4 に示す。カテゴリ間の頻度の偏りを示す指標として、0 から 1 に範囲を調節したエントロピーを下記の式を用い算出した。エントロピーが小さいほど探索が偏っていたことを示す。

$$adjusted-H = \left( \sum_{C=category} P(C) \log_2 P(C) \right) / \log_2 5$$

実験 1 では、仮説駆動条件 (.708) において、データ駆動条件 (.768) よりも有意にエントロピーが小さかった ( $t(42) = 1.850, p = .035$ )。一方、実験 2 では 2 つの条件のエントロピーに有意な差異はなかった (仮説駆動 .765, データ駆動 .769;  $t(39) = 0.102, p = .460$ )。データ駆動条件では、2 つの実験でエントロピーの値がほぼ同等であったことを考慮に入れると、仮説駆動条件の参加者は、自身の仮説が支持されなかった実験 2 においてデータ空間の探索範囲の偏りを減らし、広い範囲を探索するようになったことが示される。

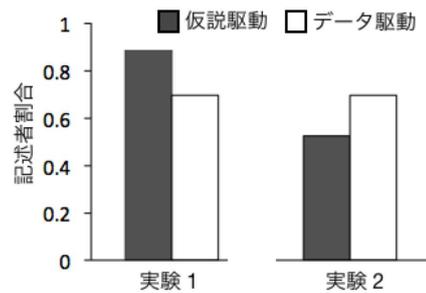


図 5 思考量の記述を行なった参加者割合。

### 2.2.4 自由記述

実験者により、導出されやすい理由として設定された「新システムを使用するとよく考えて問題を解いていた」に関連する記述を行なった参加者の人数を数えた。具体的には、「よく考えた」「思考量が増えた」のように、思考量に関する記述を行った参加者である。図 5 に記述者の割合を示す。実験 1 では仮説条件において記述者の割合が多い傾向があったが (fisher's exact test  $p = .088$ )、実験 2 では両条件に有意な差はなかった ( $p = .346$ )。

## 3. 考察

本研究は、2 つの異なる研究アプローチが、データ分析における従属変数の選択に与える影響を検討した。

### 3.1 探索の深さ

本研究ではデータ空間の探索の深さに関して、仮説駆動アプローチはデータ駆動アプローチよりも深い探索を促進するという深さ仮説を検証した。両実験で、仮説駆動条件の参加者の方が新しい変数を多く作成したという実験の結果は、深さ仮説を支持する。

さらに、仮説駆動条件の参加者が作成した変数の種類には実験間で差異が見られた。これは、彼らが 2 つの実験において異なる仮説を検証するためにデータ空間の探索を行っていたことを示す。学習効果に関する仮説が支持される実験 1 では、仮説駆動条件の参加者は学習効果の検証を短時間で打ち切っていた。ここから、彼らは主に学習効果が生じる理由を検討するために変数の作成を行っていたと考えられる。実際に、変数の作成を行った 14 名のうち 8 名は、実験者が準備した理由の検証に使用可能な、一問あたりの思考時間にあたる変数を作成していた。

一方、実験 2 では学習効果に関する仮説はデータにより支持されないが、参加者は仮説をすぐには棄却せ

ず、長い検討を続けた。この結果は、先行研究の結果と一貫している [11]。参加者は仮説を支持する証拠を見つけようと、新しく学習結果に関する変数を作成したと考えられる。これらの結果から、参加者は自身の仮説を支持する証拠が不足していると感じたときに、データ空間の深い探索を行うことが示唆される。

### 3.2 探索の広さ

データ空間の探索の広さに関しては、広さ仮説において、データ駆動アプローチは仮説駆動アプローチよりも広い探索を促進することを予測した。実験1では、データ駆動条件の参加者の方が、仮説区駆動条件よりも偏りの少ない探索を行っており、広さ仮説を支持する結果が得られた。しかし、実験2では、仮説駆動条件における探索の偏りは解消され、広さ仮説は支持されなかった。

実験2の結果は、一見、仮説が支持されない場合は仮説駆動アプローチからデータ駆動アプローチへと方略を切り替えるという Klahr の結果と一致している [12]。しかし、重要な点として、探索の深さは実験2においても仮説駆動条件の方が深かった。もし、アプローチを変更しているのであれば、探索の深さは浅くなるはずである。

我々は、実験2の仮説駆動条件の参加者は、仮説駆動アプローチを継続して用いていたと考える。仮説駆動アプローチを用いた参加者は、最初に探索を行った狭い範囲では、自身の仮説を支持するデータを見つけることができなかった。そこで、アプローチを変更するのではなく、仮説を保持したまま、データ空間の探索範囲を広げることにより、仮説を支持するデータを探そうとしたと考えられる。その結果、実験2では両条件の探索の偏りに差異がなくなったのである。

### 3.3 結果の解釈

両条件ともに参加者に同一のデータを与えたにも関わらず、実験1では、学習効果の説明に関して傾向ではあるが差異が生じた。実験で教示された2つの自宅学習アプリは、旧システムは例題を見ながら問題を解く、新システムは例題を消した状態で問題を解くというものであった。それらの特性を元に仮説をたてた場合、「新システムを使用するとよく考えて問題を解いていた」という仮説は非常には思いつきやすい。実験1で与えられたデータはこの仮説を支持するため、

仮説駆動条件の参加者はこの仮説に着目し、分析を行い、その結果記述が増加したと考えられる。

データ駆動条件の参加者では、実験1,2ともに意欲や塾に通う時間などを学習効果の理由とする、評価実験の参加者特性に着目した記述が散在した。また、同一の分析結果を得ても、同条件の他の参加者と異なる理由に帰結することもあった。これは、様々な結果を、後から統合して検討することにより、広い視点から柔軟な思考が可能であったためであると考えられる。

本実験の結果から、仮説駆動アプローチを用いると、仮説を立てる段階で一定の傾向を持ちやすく、参加者間で類似した記述が行われやすいことが示された。一方、データ駆動アプローチを用いると、分析結果に対し後から説明をつけるため、様々な解釈がなされることが示唆された。

## 4. 結論

データ駆動アプローチを用いた参加者はデータ空間を広く、浅く探索する一方、仮説駆動アプローチを用いた参加者はデータ空間を狭く、深く探索することが示された。ただし、仮説駆動アプローチを用いた場合、自身の仮説が支持されないと、仮説を支持するデータを求め、探索範囲を広げていくことが明らかとされた。今後は、結論に用いられた分析結果やその整合性、妥当性も検討する必要がある。

## 謝辞

この研究に貢献した名古屋大学情報文化学部岡村溪太君に感謝の意を表します。

## 5. 参考文献

### 文献

- [1] Van Joolingen, W. R., De Jong, T., (1991) "Supporting hypothesis generation by learners exploring an interactive computer simulation", *Instructional Science*, Vol. 20, No. 5-6, pp. 389-404.
- [2] Van Joolingen, W. R., De Jong, T., (1997) "An extended dual search space model of scientific discovery learning", *Instructional Science*, Vol. 25, No. 5, pp. 307-346.
- [3] Klahr, D., Dunbar, K., (1988) "Dual space search during scientific reasoning", *Cognitive Science*, Vol. 12, No. 1, pp. 1-48.
- [4] Shute, V. J., Glaser, R., (1990) "A large-scale evaluation of an intelligent discovery world: *Smithtown*", *Interactive Learning Environments*, Vol. 1, No. 1, pp. 51-77.
- [5] Chen, Z., Klahr, D., (1999) "All other things being equal: Acquisition and transfer of the control of variables strategy", *Child Development*, Vol. 70, No. 5, pp. 1098-1120

- [6] Grolemond, G., Wickham, H., (2014) “A cognitive interpretation of data analysis”, *International Statistical Review*, Vol. 82, No. 2, pp. 184-204.
- [7] Jolaoso, S., Burtner, R., Endert, A., (2015) “Toward a deeper understanding of data analysis, sensemaking, and signature discovery”, *Human-Computer Interaction*, pp. 463-478.
- [8] Kell, D. B., Oliver, S. G., (2004) “Here is the evidence, now what is the hypothesis? the complementary roles of inductive and hypothesis-driven science in the post-genomic era”, *Bioessays*, Vol. 26, No. 1, pp. 99-105.
- [9] Simon, H. A., Lea, G., (1974) “Problem solving and rule induction: A unified view”, In L. W. Gregg (Ed.), *Knowledge and cognition* (pp. 105-127) Hillsdale, NJ: Erlbaum.
- [10] De Mast, J., Trip, A., (2007) “Exploratory data analysis in quality-improvement projects”, *Journal of Quality Technology*, Vol. 39, No. 4, pp. 301-311.
- [11] Mason, L., (2001) “Responses to anomalous data on controversial topics and theory change”, *Learning and Instruction*, Vol. 11, No. 6, pp. 453-483.
- [12] Klahr, D., (2000) *Exploring science: The cognition and development of discovery processes*, Cambridge, MA, The MIT Press.