

OECD加盟国の生徒の学習到達度

アンドレア・シュライシャー*

2003年1月27日

於：東京大学大学院教育学研究科第一会議室

金子：国際的な学習到達度評価の調査で、非常に有名になりましたシュライシャーさんにお話ししていただけます。皆さん堪能な方ばかりだと思いますし、プライドがあるので今日は通訳は多分いらないと思います。スライドが抜けている部分があったら、私が時々補足するというような形でやらせていただこうと思います。それでは、シュライシャーさん。

1 はじめに

シュライシャー氏：ご紹介ありがとうございます。そして、お招きいただき、ありがとうございます。日本でPISAについてお話できることは、いつでも喜ばしいことです。私はドイツ出身ですが、ドイツでその結果を説明するのは非常に難しい。ドイツは、成績があまり芳しくなかったのです。しかし、日本の生徒は世界でも最高レベルに達しておりますので、結果について説明するのは気が楽です。

今日お話することは、基本的には準備してきましたが、疑問点があれば、いつでも遠慮なくそのつど質問してください。まず、テストの文脈について、手短にお話したいと思います。次に、私たちの調査の結果を見ていきましょう。とりわけ、日本に注目します。国際的に比較すると、日本はどうか。

第3に、より興味深いところなのですが、PISAで実際にできることを見ていきます。すなわち、私たちがPISAを実施した主たる理由です。私たちは、

さまざまな国における諸問題を明らかにしたいのですが、それだけでなく、現実的な到達目標はなんなのか、各国はそれぞれの条件下でどのようなことを達成できるのかということについても明らかにしたいと考えています。基本的には、他の国々にとって目標となる、最高の成果を出している国を設定することとなります。

そして最後に、とても重要な問題なのですが、PISAが基づいている方法論についてお話しします。調査結果を分析する方法、相対的な長所と短所を識別し説明する方法です。これがPISAの重要なカギとなっています。では、4つのポイントについてお話ししながら、私たちの研究の基盤となっている方法論についてお話ししたいと思います。

2 PISAの概略

2.1 PISAにおける四つの挑戦

PISAには4つのパートがあります。

第一に、学校の本質にかかわる考え方です。たとえば、どこで教育を行うかとか、何歳で学校に通いはじめるかといったことは国によって異なります。解決しなければならなかった最も根本的な問題のひとつは、公正に比較するにはどうすればいいかということでした。私たちの見解では、ある特定の学年で子どもたちを評価することはできません。例えば、もし第10学年の生徒を選んだとしても、国によって異なる年齢で入学し通学し、発達していきます。各OECD加盟国にはそれぞれ異なるタイプの制度があるのですから、そ

* OECD-PISAプログラム責任者

のようなやり方で結果を比較することはできないのです。そこで、私たちは、15歳という年齢集団で比較することを選択しました。これは全く恣意的な選択ですが、ほぼすべての生徒が、15歳という年齢で、人生の重要な段階を迎えます。もちろん、例えばフランスのような国では子どもたちは3歳で学校に通いはじめますし、またフィンランドのような国々では、7歳から学校に通いはじめます。しかし、それは政策の選択で、各国が決めることです。私たちが検討しようとしているのは、その結果なのです。評価の最低基準、すなわち年齢を標準化したわけですが、対象集団や、比較のための基準を定めるのは非常に難しいことです。

第二に、公正に比較するために、枠組みを開発しなければなりません。これは大変テクニカルな問題のようですが、非常に難しい問題でもあります。生徒にとって重要な知識やスキルとは何なのか。それらをどのように定義するのか。そして、それらの定義をどのように操作化するのか。国や文化、そして言語によらず妥当な方法で行うには、どうしたらよいのか。これは、とても困難な作業で、概念的であるとともに、技術的な作業でもあります。サイコメトリックな作業もかなり含まれています。

概念的な作業というのは、生徒にとって重要と考えられている知識やスキルについて、各国間での合意を得ることにかかわっています。これはPISAで最も困難な作業でしたが、非常に簡単だと思っていました。1つの方針しか考えていなかったほどです。しかし、これこそがPISAの最も難しい部分だったのです。各国から合意を調達するのに、数年もの歳月がかかりました。広い意味でのスキルについて申し上げているわけではありません。数学が重要であること、問題解決が重要なことは皆が分かっていますし、このレベルにおいてはみな同意するのですから。PISAは、実際にそれらを操作化した最初の国際的な研究なのです。

最初に行ったのは、科学的な概念枠組みの開発でした。これまで、国際的な研究は、多くの試験項目を集約しただけでした。私もかかわっていたので、よく知っております。多くの試験項目を集め、それら全てを混ぜ合わせ、ちょっとした内容のものを作って、こう言うのです。「テストができた」。こんなやり方では、PISAのような研究では絶対に上手くいかないでしょう。PISAは政治的なプロセスだからです。あらゆる国が懸念していたのは、公正で妥当な方法によっ

て比較されているかどうかということです。誰もがこのことに注目していました。だから、私たちはこれまでとは異なる方法を選び、科学的な根拠を開発しなければならなかったのです。例えば、ある試験項目を選ぶ際には、私たちはその項目を何となくよさそうで妥当に思えるからという理由では選ぶことはできません。私たちは、テスト項目の構造的な特徴が何であり、なぜ問題を難しくしたり易くしたりしたのかを説明できなければならないのです。30%の生徒しか正解できなかったからといって、その問題が難しいとみなすことはできません。まずはその問題に込められた課題を認識しなければならず、その上で、これは簡単にしよう、これは難しくしようと考えていくのです。

金子：PISAの話をして。PISAというのはOECDで行なった国際比較調査ですが、一つは国を超えて比較します。その年齢の点では15歳ということです。15歳の学力比較。これは、教育システムには違いがありますが、15歳という年齢を狙うということです。

ただ、その調査する内容ですが、これまで学力調査は国際的にいろんなものがありました。しかし、必ずしも問題の内容自体についてシステマチックに検討したわけではありません。それで、PISAで一番難しかったのは、内容を意味のある内容、科学的に根拠のある内容にするということです。そのためにどのような試験項目がどのような構造をもっているのか、あるいはなぜ重要なのかといったことについていろいろと議論をしたわけです。各国ともに、この試験は、政治的な意味合いをもっているのも、この部分が一番難しかったということです。

シュライシャー氏：3つ目の問題は、すでに申し上げたように、私たちは調査結果の分析をしたかったわけですが、私たちに影響を及ぼす重要な要因、個人的、構造的あるいはシステムの要因とは何かということでした。そのためには、背景的数据を生徒、学校、教育システムから集めなくてはなりません。これも非常に簡単なようですが、私たちではどうしようもない問題に直面したのです。例えば、私たちは、教室内の構造的なプロセスについて知りたかったのですが、これは無理でした。教師に長い質問紙を配ることは容易ですし、教師たちがそのアンケートに正直に答えていると信じてもいいのです。しかし、教師たちが実践していると認識していることと、彼らの教室での実践とが一致しているかどうかを、どうやって知ることがで

きるのでしょうか。たくさんの分析結果が示しておりますように、自分でやっていると思っていることと実際にやっていることとの間には大きな不一致があるのです。こうして、質問紙調査を用いることができないという方法論的な困難に直面したのです。私たちは、ビデオを用いて研究をすすめています、この領域はPISAの主要なチャレンジの1つであり、依然として解決できていない部分も残っています。

金子：調査の時に非常に重要なのは、結果としての学力にどのような要因が効いているかということです。例えば学校とか、子どものデータも集めているわけですが、一番難しかったのは教師の教育の仕方に対するデータを入れることです。それは教師に質問表を配って、調査をすることもできるわけですが、それは実際に教室で行なっていることと必ずしも一緒ではありません。ビデオを使ったような調査もやっていますけれども、その教育のプラクティスがどのような要因として働いているかというところが最も難しいところです。

シュライシャー氏：PISAを開発する中で発展した研究分野に、各国が到達目標を議論するための共通分母の確立があります。当初はそれが重要な分野になるとは私たちは想像していませんでした。この分野は最初の段階から存在していたわけではなく、研究を進めていくうちに予期せず発展したものなのです。

現在、PISAは、各国の協議者が情報交換の場で到達目標を設定する際、主要な根拠として用いられています。このことは、私たちが当初計画していたことではなかったのですが、今では、よりシステムティックなやり方でPISAの調査デザインに組み込まれています。これは、PISAの最も重要な機能の1つへと発展しており、各国に利用されています。例えば、各国は、教育課程をPISAと照らし合わせて、PISAにどのように対応しており、共通点が何で、相違点が何なのかということを調べています。このように、PISAは、各国がそれぞれの教育目標や到達目標などを議論するためのひとつの枠組みとして使われています。

金子：国際的な学力比較というのは非常に難しいわけです。しかし、逆にこういった努力をしていったためにPISAの試験が国際的な学力の公約数、標準みたいなものになって、各国が自分の国の問題を考えるときに、PISAを一つのスタンダードとして考えるということもできています。

2.2 PISAの位置づけ

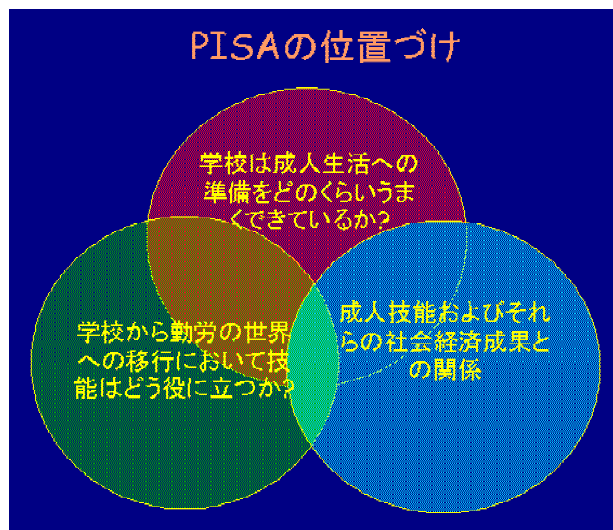
シュライシャー氏：もう一度、PISAの研究対象に話を戻しましょう。PISAが本当に回答しようとしている問いは、生徒たちが世間や生活や仕事など学校を卒業した後のあらゆることに対して、どの程度準備がされているのかということでしょう。これが、私たちが取り組もうとしている根本的な問題です。

しかしそれだけではありません。私たちは、学校から仕事への移行といったテーマにも目を向けはじめています。生徒たちが学校を離れた後、どのようなことが起こるのでしょうか。どのような選択をするのでしょうか。態度はどうでしょうか。学校で獲得したスキルは、継続教育further educationや仕事に対する態度とどのように結びついているのでしょうか。

私たちは現在、PISAを拡張して、縦断的調査〔訳注：同一対象者を長期間継続して追跡する調査方法〕をしています。この調査は全ての国を網羅しているわけではありませんが、こうした調査は非常に重要であると考えています。私たちは2つの国でこの方法論を試しました。そして、今では比べ物にならないほど多くのサンプルを採るまでになっています。通常、PISAのサンプル数は5000から50000人の生徒ですが、この種の研究を行うには、最初の段階で、かなりの生徒数が必要となります。さらに、質問紙調査を毎年行うための追跡調査方法が必要となります。非常に労力を要するのですが、この調査によって私たちはより深く洞察することができるでしょうし、このような調査によってのみ、因果関係についての問いに足を踏み入れることができるでしょう。クロスセクショナルな調査では、因果関係に堪えて回答することは決してできないでしょう。継続的なサーベイによってのみ、各生徒を追跡することができるのです。これが、私たちが辿りついた答えです。

金子：PISAは、学校が、生涯にわたりどのような学力上の準備をしているのか、成功しているのかということにすごく大きく焦点を当てているわけです。もうちょっと具体的にいうと、学校から勤労へ、職場・職業への移行に具体的にどのような役割を果たしているのかということ、いくつかの面でみていくことが必要になってきます。そこに書いてあるような項目が焦点になってくるわけですが、それを分析するために、二つの国で追跡調査もやっているということです。そ

の追跡調査をやらないと具体的に学力がどのような社会行動に結びつくのか、というのは分析できない。



シュライシャー氏：赤い丸のところがPISAを示しています。この青い丸や、学校から職業への移行を示す緑色の部分は、私たちが近年に取り組みはじめているものです。そして、次の青い丸ですが、参加国と討論するなかで、各国の人的資本の蓄積を査定する方法を開発しようとしています。私たちはその際、同時にスキルと職業の関係についての問いも検討することができます。例えば、産業におけるスキルの需要と供給との関係のような問いです。まだその理論的な枠組みを開発しはじめたばかりですので、この場で私がお話できることは何もありません。しかし、これら3つの丸（赤、青、緑の丸）全ては相互に深く関係しているので、それらを組み合わせることで、一国におけるスキルの分布状況を知ることができるでしょう。

金子：PISAの位置付けの第三の分野は、社会でどのような能力が必要とされるのか、ヒューマン・キャピタルというのは人的資本、どのようなものが必要とされるのかという分野です。これについても分析を計画しているわけですが、まだこれは計画途中であるということです。いまは具体的にお話ができないわけですが、この3つの分野で活動するというのが我々の目標です。

2.3 PISA調査が評価するスキル

シュライシャー氏：ここまでがPISAについてのお

話です。先ほど、お話ししたように、評価を操作化することは難しいのです。ここでは、OECD30カ国がスキルとして重視しているもののうち、重要な分野に注目したいと思います。

一般的に、リテラシー・スキルは、学校におけるあらゆる事柄の基礎となっていると理解されています。しかし、PISAでは、リテラシーは、読み書き能力と定義されてはいません。なぜなら、それでは、技術的なリテラシーにすぎないからです。私たちが実際に用いたのは、その背後にある概念で、“世界への扉”という意味で用いています。すなわち、情報にアクセスし、情報を処理すること、情報を結びつけたり、情報を評価したり、そして情報にもとづいて熟考するといった能力です。これが、私たちのリテラシーの定義なのです。

金子：基礎的な能力というのをどのように捉えるのかということです。まず重要なのはリテラシーですが、リテラシーというのは単に技術的な意味でいっているのではなくて、いってみれば社会に対する、あるいは世界に対する窓というか、情報を獲得するといった意味で社会に対する窓という点で重要だということです。シュライシャー氏：リテラシーといえば、これ自体は簡単に聞こえると思います。しかし実際には、日本の生徒たちが、最も困難を抱えている課題だったのです。数学と科学は1位と2位です。しかし、読解リテラシーについてはご覧の通り、評価したり、熟考したり、推論したりといった幾つかの分野に関して、日本の生徒は苦手としているようです。

第2の分野については、何と言いますか、複雑なものに対応したり、問題を解決したりする能力です。この分野は、PISA2003調査に取り入れようとしているものですが、一回目は盛り込むことができませんでした。しかし、いまでは、問題解決能力、すなわち、パターンを認識して類似性を見出し、発展させたり、問題を認識して、解決の戦略を展開したり、その戦略を評価するといった能力を、査定することができます。こうして、リテラシーという概念に新たな要素が持ち込まれました。

次にお話する分野は、私たちがいまだに奮闘している分野です。上述の2つの分野に関してはかなりカバーできたと思いますし、そのための有効な方法も手に入れています。しかし、対人関係能力については、まだ妥当な評価方法がありません。おそらくこの対人関

係能力については、国家や特定の文化環境の中では評価しえるものなのかもしれませんが、文化的背景を超えて評価するのは非常に困難です。なぜなら、基本的には、人々は、文化や国によって異なる方法で交流しているからです。しかし、そうは言うものの、人生に対してどれだけ準備ができているかを左右する重要な能力のひとつだと皆認識しています。だから、将来この評価ができるように、現在私たちは対人関係能力という分野をPISAの枠組みの中で発展させているのです。この分野には、コミュニケーションだけではなく、グループに参加する能力やグループで何らかの機能を果たす能力、そして一緒に作業する能力など、他人の役割も果たすための全ての能力、基本的には他人と一緒に暮し、働く能力が含まれているのです。

金子：リテラシーの次は複雑さへの対応というか問題解決技能といったものですが、この二つはまだ比較的計測しやすいわけです。第三の我々が重要だと思うのは人間関係の技能で、それは人とコミュニケーションしながら、人と協力して仕事を進めていくことです。これは非常に重要だと分かるわけですが、どうやって計測するのか、それは非常に難しく、これは我々の一つの課題になっています。

シュライシャー氏：第3の分野は対人的なinterpersonalスキルでしたが、最後の分野は個人内のintrapersonalスキルです。これらは、やる気やメタ認知など、明確に定義づけするには非常に困難ですが、私たちにとって大変重要な能力です。例えば学習戦略がそれに当たるわけですが、PISA2000で評価しようとしたのは、生徒の学習への取り組み方でした。すなわち、学習方法や、自己への配慮や関与の仕方です。結果をお見せするときに、またこの続きをお話いたします。これらは、PISAテストにおいて生徒の成功を左右する重大な規定要因ですが、この分野に関しても、手段がほとんどないため、現段階ではこの分野を総合的に評価できるとは言えません。

金子：今まで述べた最初の三つは、人間とその他者との関係に関わるものです。最後に重要なのは人間の中で認知的にどのようなスキルがあるのかという、学ぶ力とか、そこに書いてあるようなことです。要するに個人の中の認識過程というものがどのように働いているのか、それは非常に重要ですが、これについてもなかなか計測するのは難しい。

シュライシャー氏：以上がスキルのフレームワークで

す。私たちはこれをどのようにして作成したのでしょうか。PISAの基本的な取り組み方は、既存の方法論を用いて、現時点で可能なことから始めるというものです。しかし、それと同時に、私たちは理論を発展させるために、世界中の専門家と共同研究もしています。ただ単に私たちが評価できるものだけを対象としているのではなく、評価できていない未開発分野にも取り組んでいます。私たちはOECDの国々に約300人の専門家を擁しており、彼らは、新しい分野の評価方法を発展させ、また、それを国を超えて信頼性のある方法とするための研究をしています。

このように、私たちは、この分野を発展させるために大変精力的に取り組んでいるわけですが、なぜこのような取り組みが必要なのでしょう？ 多くの専門家や政策立案者が、評価によって自分たちの視野が狭くなるのではないかと懸念しているからです。例えば、もし私たちがテストしやすいものだけ、数学的知識の再現といった、何でも非常に簡単にテストできるものだけに限定して評価してしまえば、人々はその分野だけに注目し、評価しやすい分野だけで、教育システムの成否を判断することでしょう。このようなことがPISAの中で起こらないように、多くの資源を投入して、評価を拡張し、教育の成功を評価するために統合すべき新たなスキル領域にも目を向けているのです。

金子：PISAというのは進行中のプログラムですが、今でも世界中で300人くらいの専門家が協力して仕事を進めているわけです。それはなぜかという、これまでのテストにも一定の意味があったわけですが、計測し得るものだけに限っているとどうしても視野が狭くなる。基本的に何が 필요한のか、能力として何を考えるべきなのか、何を視野に置くべきなのかということを常に考えて、視野を拡大していくということが、このプロジェクトで非常に重要な点です。

2.4 調査サイクル

シュライシャー氏：それではPISAはどのように行われるのか。私たちはPISAを発展させ、3年に1回の査定を実施することとしています。なぜ、そのような方法をとっているのでしょうか。私たちは、PISAを一回限りの調査として企画したわけではなく、長期的で一貫したプロジェクトとして考えています。

そのように考えるのは、私の多少個人的な経験が関

係しています。以前、IEA〔国際教育到達度評価学会〕という別の組織で働いていた頃、私は各国の学力評価を行っていました。その際、常にひとつの査定方法を用いておりました。結果を出してみると、成績が芳しくない国が分かるのですが、政策立案者らは後になって「問題を解決するために、昨年、あらゆる地域に報告書を配布しました。だから、これ以上心配することはない」と言うのです。ですから、私たちがOECDの国々でこのPISAを始めた際、私たちは、同じ過ちは二度と繰り返さないと誓ったのです。私たちは、全ての国々に対し長期的なプロジェクトであることを伝えました。私たちは頭がおかしくなったと思われるようですが、PISAを真剣に捉えてもらうには、これが唯一の方法だと考えたのです。

こうして私たちは3年間のサイクルになるようにプロジェクトを開発しました。3年毎というのは、適当な期間ですし、私たちの資源の面からもちょうど良かった。順位もモニターできます。幾つかの指標を用いることで、各国の順位をほぼ予測できるので、現在の国がうまくいってどの国がうまくいっていないのかということを私たちは知ることができます。しかし、どの国が進歩しつつあり、どの国が低下しつつあるのかということを調べるのも、非常に興味深いことです。これがひとつの特徴となっております。

金子：それからもう一つPISAの特徴は、3年ごとに調査をやるということで、これはIEA、国際協力協会の前、学力調査というのを国際的にやられていたわけですが、その経験からいうと、一回やってそれでおしまいということになると、結局その時点で一定の結果が出れば、それ以上何も意味がないということになるわけです。しかし、それではやはりいけないだろう。こういった調査をやるとどの国がいいのか悪いのかというのが大体分かるわけですが、それで終わるのではなくて、長期的に国々がどのような変化をしているのかということに我々は注意を向けたい。

2.5 PISAテストの特徴

シュライシャー氏：リテラシーという言葉についても一度話を戻しましょう。それは、情報にアクセスし、処理し、結合し、評価する能力でした。それはつまり、分析、比較、対照、そして評価できるということです。私たちはまた、生徒が大変創造的に知識を使えること

を期待しています。皆さんがご存知かどうかは分かりませんが、PISAの課題の多くは非常にむずかしいです。異なる分野からの情報を結合しなければならないのですから。しかし、それこそ私たちが求めたことだったのです。

私たちが生徒に期待していたのは、実生活において、彼らが持っている知識を活用できることなのです。時々なされるPISAへの批判は、まずもって「これは実際、数学的課題というよりは自然科学的課題です。PISAではこれを数学的なものとして分類していますが、それは間違いではないか？」というものでした。これに対して、私たちはいつも、「実生活とはそういうもののなのです」と答えます。実生活では、生徒たちはどの教科を問題解決に適用すべきかということを教えられるわけではないのです。生徒たちは異なった分野の知識を統合しなければならないですし、適切な科学的専門分野を応用する必要もあるでしょう。ですから、私たちはさまざまな事柄を概念レベルで組み合わせ、それらを選り分けるのです。生徒たちに受けさせるテストの中で、私たちは意図的に教科領域を混合するようにしているのです。

最後になりますが、私たちは生徒たちに課題を解くことだけを期待しているわけではありません。同時に彼らが自身の解答を他者に伝達できることも期待しているのです。ある国々の生徒たちは常に正しい解答を出すのですが、その解答の説明を求めると、彼らにとってはそれが非常に困難であったりするのです。私たちは、そのような能力を評価する部分をPISAの中に作りました。これがPISAの特徴となっております。

金子：PISAで考えているリテラシーをもう一回考えましょう。物事を分析したり比較したりする能力、あるいは創造的に考えるための能力、そのためのツール、その情報を活用する、知識を実際の場面で応用するというのは、学科で分類されるような知識ではなくて、実際に役に立つための知識というものが、実際の形でどのように役に立つのかということを焦点にしました。

それから、考えやアイデアを意思疎通するというのは、例えば国によっては、答えは正しく出るのだけどその説明ができないといったこともある、そういった意味できちんと説明できるといったことも重要である。こういった多面的な次元でリテラシーを考える。

シュライシャー氏：リテラシーは3つの要素を持っています。第1の要素はリテラシーに関する従来の認識である読解リテラシーです。しかし私たちはこれに数学的リテラシーも科学的リテラシーと同様に統合したいと考えています。ですから、数学、科学、そして読解の全てがPISAのリテラシーなのです。

では私たちはなぜそれをリテラシーと呼ぶのでしょうか。それは、数学と科学がもはや選り抜かれた人のための教科ではなく、世の中に出て行ったとき誰もが身につけている基礎知識となるべきものだという信念が、PISAの底流にあるからです。ですから私たちは生徒たちに多くの知識を再現することを求めたわけではありません。実際、ほとんどのPISAの課題は、生徒たちが必要とする事実的な知識は全て与えています。全ての情報を一方のページに載せ、課題を他のページに載せました。

課題に込められていたもの、言い換えれば私たちがテストで評価したいのは、次のことです。すなわち、生徒たちが科学的・数学的に考え、証拠に基づいた結論を導きだし、一般に信じられていることと証拠に基づいた結論とを明確に区別できる能力です。私たちがPISAで評価しなかったのは、こうしたことであり、知識の再現ではないのです。だから、私たちは、テストの中で生徒たちが必要とする情報を全て与えたのです。

金子：それから、もう一つリテラシーの定義です。単に文字を読むのではなくて、数学的あるいは自然科学的な能力というものを非常に重視したい。それはなぜかといえば、これからの社会では自然科学とか数学的な能力が非常に重要だからです。PISAの試験でも事実に関するものは、ごく一部にすぎなくて、むしろ一定の事実から何を類推するのかといった、数学的あるいは論理的な思考法を抽出するものを重視していきます。

3 分析結果

3.1 各国の習熟レベルの分布状況

シュライシャー氏：では、結果の報告に移りたいと思います。結果をまとめた日本語の短縮版もあるのですが、全てを記述しているわけではありません。そこで、

他の部分については私が説明いたします。まずは結果を見てみましょう。

PISAにおける重要な方法論的概念の1つに、私たちが有効性efficiencyのレベルと呼んでいるものがありますが、これは本当に重要な概念だと考えています。普通、学校では、生徒たちは成績を告げられます。たとえば、通知表でAを獲得したとします。日本でどのようなかは知りませんが、生徒たちは何らかの評価を受け、そして彼らはこう考えるのです。「この成績に実際どんな意味があるのだろうか」と。生徒たちは何ができて、何ができないのでしょうか。生徒たちは試験の準備をします。しかし、彼らは、成績と、実際に身についた知識やスキルとを結びつけることができないのです。国際テストでも、私たちは同じ問題を抱えています。点数を出して、この生徒は550点取っていますとかその生徒は330点でしたというようなことを言うのは非常に簡単なことです。もしくは、この国は成績が良くて、この国は成績が悪いといったようなことを言うのも簡単です。しかし、それらの点数と意味とを結びつけることは極めて難しい課題です。

そのために、かなり高度な方法論を適用しなければならなかったのです。まず、尺度を作らなければいけませんでした。生徒や試験項目を位置づけるための概念的な基準が必要なのです。すでに確立された方法論でしたから、いずれの理論でも、十分でした。もっと難しかったのは、試験項目に込められた課題を実際に尺度上に位置づけること、そして、試験項目がその尺度上で異なるレベルに位置する理由を説明することでした。

金子：もう一つPISAの試験の結果を見ると非常に重要なのは、PISAでは五つのリテラシーの習熟度レベルというのを考えています。そこで問題になるのは、点数というものをどのように解釈するのか、どういう意味をもっているのか、特に国際的な比較をするときにどのような意味をもっているのかということです。そのためにもちろん基礎的にはいくつかの能力のディメンションを作って、その中でいくつかのスケール、それぞれのスケールでもって能力を測るわけですが、それぞれのスケールの一種のプロフィールをみつけるわけです。

シュライシャー氏：習熟レベルproficiency levelという概念は、非常に重要です。これもまた政策協議において非常に役立つものです。なぜなら、私たちが答え

を出そうとしている重要な問いとは、相対的に誰が優れているのかということだけでなく、到達という観点から見たとき、十分優秀とは何であり不十分とは何なのかということだからです。そのような判断を下すためには、私たちは単に相対的な点数のみを用いるだけではなく、生徒の到達度についての絶対的な評価基準を持たなくてはなりません。得点から、生徒たちが身につけているスキルや能力を把握できなければならないのです。これが、PISAの主要な研究課題でした。もちろん、非常に大変な作業ではありましたが、結果をお伝えする上で非常に役立つものでした。

金子：試験をすれば点数が出るわけですが、その点数から解釈するときにそれがどのような意味をもっているのかということをは何かの形で表現できるような手段をしたい。そういう意味では、何らかの絶対的な基準というものを作りたい、それをPISAでやったわけです。

シュライシャー氏：ですから、例えば、世界中でもしくはOECDの国々の中で習熟レベル5の生徒数が何人いるかということが分かります。ちなみに、このレベル5というのは、十分な知識を身につけて利用することができ、自身の仮説を発展させることができ、新たな知識を生み出せるレベルで、つまり情報の処理から情報の生産ないし情報の評価に進んだレベルにあるということです。OECD各国の平均は10%に上ります。ここにある資料では、ドイツで9%、日本で10%です。お分かりのように、10%に至っていない国もありますが、10%というのは平均ということです。イギリスではなんと16%ですし、アメリカでは12%で、非

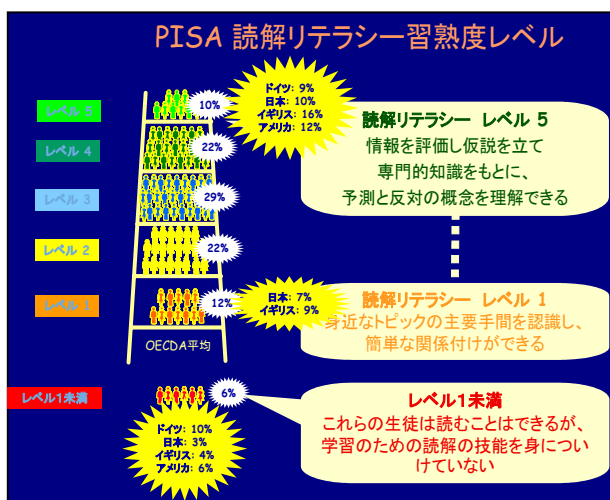
常に驚くべきことです。合衆国は全体としては、教育がうまく機能していないのですが、一方で習熟度トップレベルが、非常に高い割合を占めています。

金子：見たとおり、こういう定義をして、レベル5というのは定義すると、あのような数値、それぞれの国でレベル5に達する人の割合が、ああいう風になるということです。アメリカ、一般的には、平均的というのはあまりよくないのですが、一番上のところは比較的多いです。

シュライシャー氏：では次に、スケールの逆側つまり習熟度の低い方でも同じ事が考えられるはずですが。習熟度レベル1の生徒というのは、技術的に読むことはできるが、知識を広げるツールとして読解、数学的、科学的リテラシーを使えない生徒たちです。そのレベルは、私たちが危険なグループとみなし始めるレベルであり、レベル1よりさらに下の生徒などは、学校でやっていることに全くついていけないレベルの生徒であるといえます。そのような生徒は十分な読解リテラシーがないということなのです。例えば、数学を学んだり、教科書を理解するという能力もないということです。レベル1とレベル0と私たちが呼んでいる生徒たちは、将来における成功に大きな困難があるレベルと分類されています。

このことは、政策立案者たちにとっては、非常に重要なことなのです。もし、相対的な点数しか分らないのであれば、「これらのレベルの生徒というのは、全体の下位20ないし10%の生徒たちです」としか言えません。それで十分だという人もいでしょう。しかし、この方法論を用いることで、彼らが実際に抱えている欠陥を認識できるようになります。興味深いことに、日本は危機的な生徒の割合が最も少ない国の1つなのです。日本のシステムは、この決定的な境界周辺にある生徒たちを何とか引き上げることで、彼らが独力で学び、後の人生で自身の知識を広げられるようにしているわけです。それは、教育の、非常に重要な機能なのです。

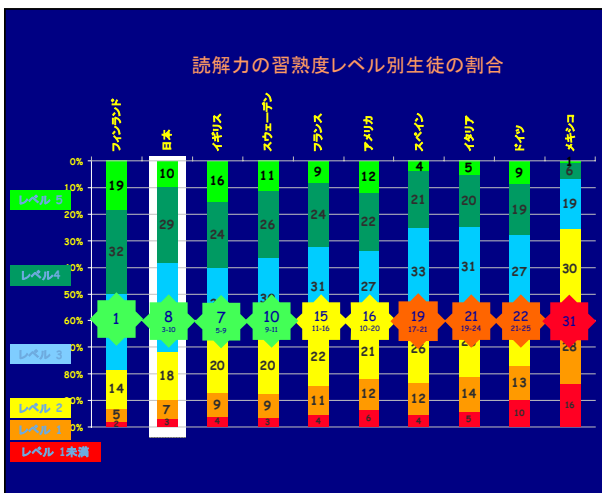
金子：今度は低いほうを考えてみると、レベル1というのは、そこに書いてあるような形で定義すると、そのような割合に大体なってきます。さらに、それに達しないレベル1未満、レベル0とよんでいますけれども、そういった人たちもいるわけです。相対的な点数であると、絶対的にそれはどういう意味をもつのかということはあまり分からないわけですが、こういうふ



うに定義したうえで、レベル0の生徒がいるということ自体、非常に大きな政策的な問題です。

シュライシャー氏：もちろん、ほとんどの生徒は、その間のレベルのどこかに入っています。次に、それをより詳細に検討することにしましょう。ご覧のように、さまざまな習熟レベルについての結果は記述できます。今ご覧になっているものですね。トップレベルにおいては、日本とドイツはほぼ等しいと言えます。しかし、次の点ではっきりと区別できます。日本には習熟度の低い生徒はほとんどいないのですが、ドイツには低い習熟度の学生が結構いるのです。

結果を要約したものを評価することもできます。けれども、これはメディアが好んですることですね。さらに各国に順位をつけ、このように並べ替えることもできます。しかし、調査結果を分析するという点では、それはあまり意味のあることではないし、これはとても重要なことですが、各国のランクを正確に把握することは決してできません。なぜなら、その数値はサンプルに基づいているのですから。各国が占める順位の範囲しか正確には分からないのです。

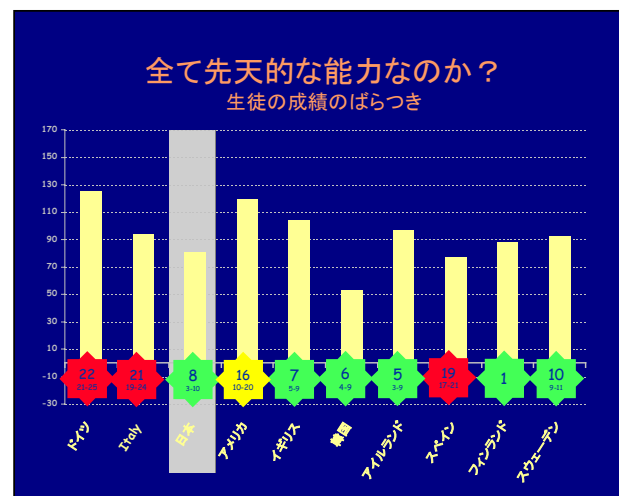


金子：先程の5段階、プラス6段階の分布を各国についてみるとそういうことになるので、日本とドイツというのは、トップのほうはそんなに変わらないけれども、非常に違うのは下の方の成績の人たちが日本では少なく、ドイツでは多いということです。ただここには簡単に国でランクづけをすることもできるわけですが、むしろ重要なのは分布だということです。

3.2 習熟度の分散を説明する

次は分析的な問いです。今までは単に科目の話でしたね。まず私は読解リテラシーの分野についてお話ししました。そして、日本は、科学で第二位、数学で第一位であると申し上げました。

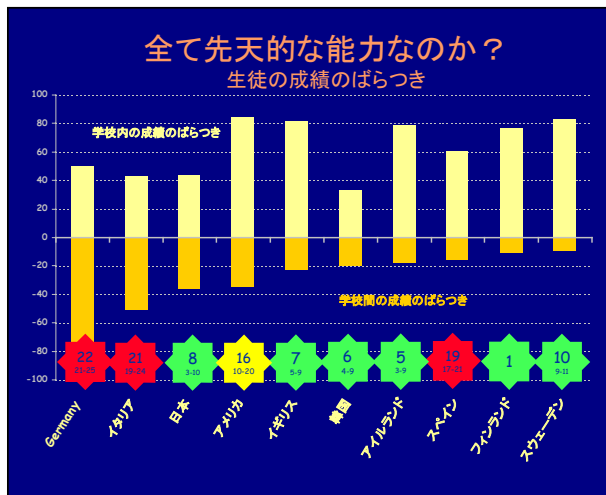
次なる疑問点としては、「最低レベルの習熟度にある生徒と最高レベルの生徒との間にある違いは何なのか」ということでしょう。結果をグラフにプロットすることができます。このバーの長さは分布を示しています。この分布の様子は、私たちが興味を持った最初の調査結果でした。この結果は、政策を考える上で、非常に興味深いものなのです。なぜなら、教育システムにかんして一般に信じられていたことが真実ではないということを、この結果は示しているからです。すなわち、得点の平均を高くしたいのであれば格差が大きくなることを受け入れなければならないし、あるいは逆に、格差を最小にしたいのであればその質は低下するという説は正しくないのです。OECD各国の中で1位のフィンランドを見てください。習熟度はナンバーワンであり、明らかに他の国より上です。そしてさらに、彼らは、その習熟度の格差を狭い範囲に抑えることにも成功しています。ですので、国のランクをその分布とともに見てみると、興味深い考察ができるのです。別の例としては、韓国は非常に習熟度が高いですし、日本の習熟度も高いです。そして、これらの国々もその格差を狭い範囲に抑えています。



金子：そのPISAのスコアで非常に面白いのは、成績の絶対値でなくてばらつきです。これは分散を示し

ているものですが、一般的にいえば、平均点を高くしようとすると、ばらつきが出る。それから、ばらつきを抑えようとすると平均点が余り高くないというように思われがちですが、実際にはそうではなくて、平均点が非常に高いところで、分散が低い国、例えばフィンランドあるいは韓国、日本のようなものもあるし、それから非常に分散が高くて平均点も余り高くないという国もある。

シュライシャー氏：では、次の疑問点としては、「この分散はどこからくるのか」ということでしょう。それは、各学校のレベルが全く異なっているからなのか、それとも、学校内で生徒の習熟度レベルが大きく異なっているからなのか。私たちのような方法を用いれば、このような問いにも答えることができるのです。なぜ、以前の研究ではそのような問いに答えることができなかったかと言いますと、それまでの研究では、各学校の数クラスのみを評価していたからなのです。私たちは対象年齢の全生徒を用いています。そのため、各学校の習熟度がどうなのかが分かりますし、そこから、学校内での分散と学校間での分散の両方に分解することができます。こうして、学校内での分散と学校間の分散を識別することができるのです。



ここで、もう1つの興味深い事柄に気づかれると思います。フィンランドやスウェーデンの結果を見てください。これらの国は概して優秀で、共にグリーンエリア〔10位以内〕に位置するわけですが、OECD諸国の平均より明らかに上であり、得点の分布も比較的狭い範囲に収まっていますし、学校間の差もほとんどありません。ですから、フィンランドでは、親はどの

学校にでも子どもを通わせることができます。ちょっと大げさかもしれませんが、どの学校に通わせても、質の高い教育を受けられる保証があるようなものです。私たちの用いたサンプルでは、そしてこの年齢のサンプルによれば、学校間の分散は10%以下なのです。

なぜこの方法論がいいのか、この方法論の強みがお分かりでしょう。もし、これを数学級だけしか対象にしていなかったら、このような分析は全くできなかったはずですよ。

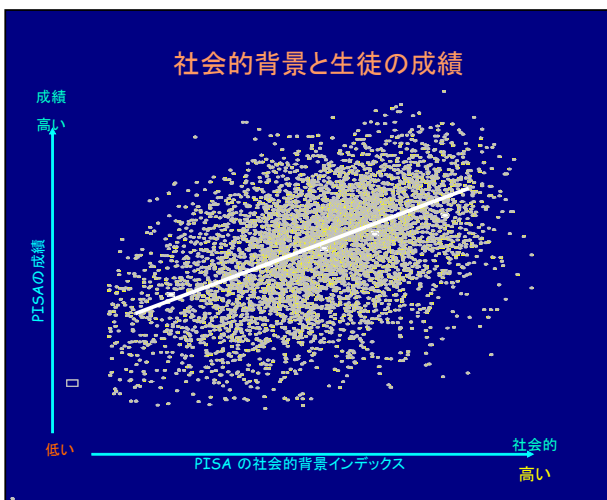
次に、スケールの逆側、習熟度が低い方を見てください。全体として達成度の低いドイツでは、分散が大きく、かつその分散のほとんどは学校間での格差によるものです。これについては、また後で、もう少し深くお話しと思いますが、これは、主に教育システムの仕組みの結果です。だから、同じような構造を持つ他の全ての国で同様の結果を見ることができます。

ドイツを弁護するために、ここで述べておきたいのですが、PISAテストはドイツの教育システムで教えられてきたことには向いていないのです。ドイツでは、知識の再現や処理managementにかんする課題は非常に良くできています。しかし、知識を創造することにかんする課題、すなわちPISAのテストで重視されている課題では、点数が低くなっています。しかし、分散は構造的な問題です。この後、より詳細に分析することにいたします。

金子：もう一つは成績のばらつき、分散はどこからきているかということです。PISAは学校の中でかなり十分なサンプルをとっているのです、学校間と学校の中での生徒のばらつき、分散を分解することができるわけです。それをやったのがこのグラフで、下の方が学校間のばらつき、上のほうが学校内のばらつきです。比較的点数が高いフィンランド、スウェーデンなどをみると、面白いのは学校の間のばらつきがむしろ少なくて、学校の中でのばらつきがあるということです。ただ、平均点が比較的低いドイツとかイタリアでは、学校の間のばらつきもかなり大きいというところが、非常に重要なところですよ。ただ、ドイツで点が比較的悪いというのは、やはりドイツの学校で教えている内容がかなり違うようで、知識を述べるテストが比較的ドイツの生徒はかなりよくできるのですが、類推とか、知識を創る、そういった点が弱いので平均点が低くなっているわけです。ただ、平均点はそういうふうな影響を受けているかもしれないけれども、実際にばらつ

きが大きいということはこのグラフで非常によく分かります。

シュライシャー氏：次の質問としては、規定要因が何なのかということでしょう。私たちの分析で重視したのは、各個人の社会的背景です。では少しいりくんで、複雑な図をご覧にいきましょう。ここに習熟度の低い学生から高い学生へとプロットしていきます。他方で、私たちは、生徒たちに彼らの社会的背景について多くの質問を行いました。父親の職業、親の収入、家族の裕福さ、社会的コミュニケーション、文化的背景などについてです。社会経済的背景に関する多くの質問をしました。そして、全ての生徒をこの図の中にプロットしていきます。これは恵まれない環境の生徒で習熟度が低い子、この生徒は、恵まれない環境の子でも習熟度は高い子、そして、この生徒はより特権階級の家庭の子です。こうして、生徒たちはそれぞれプロットされて、回帰直線に要約されるわけです。OECD各国すべての結果です。私たちは各国2000人の生徒のプロットを行い、全体で25万人を超える生徒を評価しました。



私たちの報告を見れば、多くの人が非常に失望します。なぜなら、社会的背景が習熟度に影響を及ぼす主要な要因となっているからです。私たちは皆、平等な学習機会を目指していますが、この調査結果は、その方法ではうまくいかないことを示しているのです。

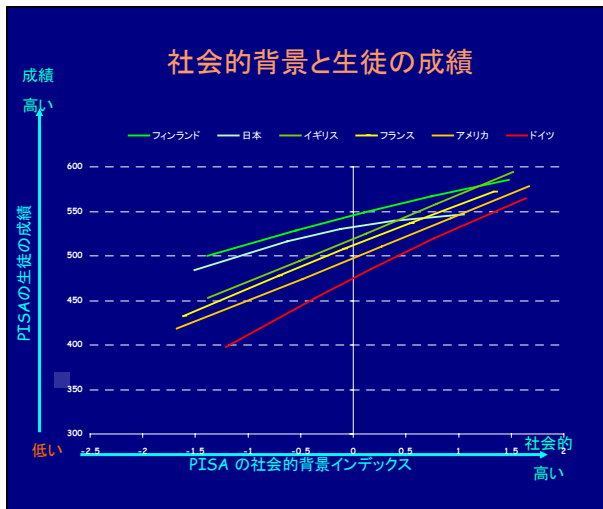
他方で、分散が大きいことに注目すると、不利な環境で育ったたくさんの生徒が成功していることが分かります。その背後には何があるのでしょうか。神が与

えたもうた平等なのか、あるいは、教育政策が何かをもたらしたのでしょうか。この問題に回答するには、全ての国で同様の現象が見られるのかを検討することです。分析のために、私たちは習熟度の尺度だけでなく、社会的背景についての尺度も必要になりました。例えば、両親の職業についてでしたら、課題を尺度上に位置づけるように、親の職業についても並べておかなければならない。それは些細なことではなく、非常に重要なことなのです。さもなければ、この種の分析は不可能となるからです。個別に質問するだけで、その情報を後から混合することはできません。フレームワークの段階から考慮されていなければならないのです。

今回は、幾つかの国を選んでお話ししましょう。フィンランドは平均が非常に高いということを知っていましたが、回帰直線は、OECD各国よりもはるかに平らです。ですから、フィンランドはその他のほとんどの国に比べて、社会的背景が習熟度に及ぼす影響を抑えることに成功していると言えるでしょう。日本についても、ほぼ同様のことが言え、日本は社会的背景の影響を調停するという観点では最良の国の1つですが、しかし、回帰係数を見ると、34.4ポイントという数字などは非常に大きい。次に、イギリスを見てください。非常に急な勾配になっています。もしあなたが恵まれた社会環境にいるなら、あなたはフィンランドの生徒と同じくらい良くできるはずですが、しかし、もしあなたがこの位置に生まれたならば、統計的に言えば、あなたはOECD諸国の平均を大きく下回るということになるでしょう。フランスも、多少習熟度の平均が低いということ以外はイギリスと同じような様子です。アメリカ合衆国はさらにもう少しレベルが下がります。そしてここで例として挙げているドイツは、他にも同じような国がありますが、社会的背景が生徒の成功の主要な要因となっております。

繰り返しますが、もしあなたがこの位置に生まれたなら、不運です。これは実に興味深い図です。この図が示していることは、最高の習熟度の生徒にとっては、教育システムによる違いは、それほど大きくないということです。教育システムに違いをもたらしているのは、習熟度の低い生徒を処遇し解決する能力です。これは、この分析で明らかになった非常に重要な部分です。

金子：成績の差がどのようなところから出ているかと



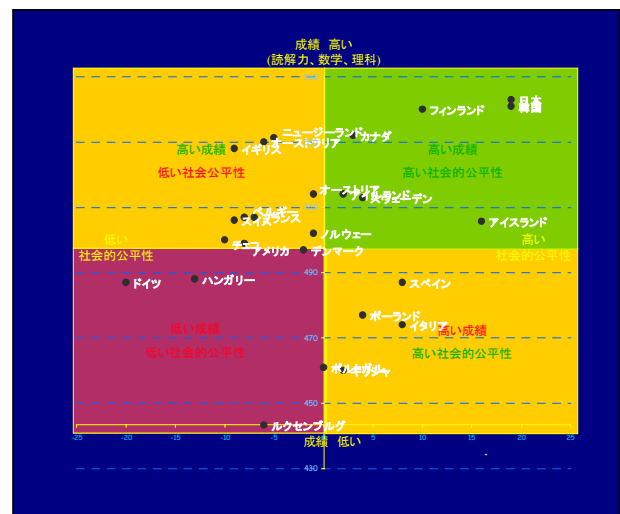
ということです。両親の社会的な背景というのが一つの重要な要因になるわけです。こういうふうに散布図を描いてみるとこのようになって、かなり有効な関連がみられるわけですが、やはりその散らばりがかなり大きい。各国別に今度は相関図、相関関係等、その回帰直線を描いてみると、このようになるわけです。それでフィンランド、日本は、比較的社会的な背景との関連が薄い。それから、イギリス、フランス、アメリカ、ドイツはかなり社会的な背景の影響を受けている。これを通じて非常に重要なのは、むしろアチーブメントの低い層の問題だということがよく分かります。

3.3 達成可能な目標

シュライシャー氏: これまで、習熟度という観点から、自分たちがどこに位置し、その他の異なる国々がどこに位置するのを見えてきました。次なる問いは「現実的に達成できる目標は何なのか」ということです。そこで、あらゆる国を再び尺度上にプロットします。全ての教科（読解、数学、科学など）を混合した尺度です。全ての国をこの尺度上に載せると、ご覧の通り、日本は一位になります。これは各国の平均習熟度です。

しかし、ここで、もうひとつの軸を加えてみましょう。政策上の非常に重要な問題点は、教育機会の質にかかわるものです。ひとつの次元だけで分析するのではなく、機会の質にかかわるもうひとつの軸を加えたいのです。多くの勝者や敗者がいて、社会的な差異が大きな社会なのか、あるいはより公平な社会なのかと

いう軸です。そこで、社会的不公平から社会的公平までの第2の尺度をつくりました。ここで、2つの軸を用いるのは大変に重要なことなのです。基本的にこの第2軸は、社会的背景が生徒の習熟度に及ぼす影響の大きさを示すものです。皆さんは理想としては、その社会的背景の影響がなく、生徒が彼らの本当の能力によってのみ測定される状況を期待するでしょう。理想的な教育システムではこれが可能だと思います。この図で、左に行けば行くほど、状況はますます悪くなり、右に行けばいくほど、状況は良くなります。誰もこの4角形の中にいたいと思っていますし、誰もこちらにいたいとは思っていません。そしてこれら2つについては多くの議論がなされています。ここにその典型的なものを挙げます。先に私がお見せした回帰直線の傾きから、すでに想像なさっていると思いますが、これはイギリスです。全体として高度な習熟度ですが同時に大きな社会格差があります。イタリアは、イギリスとは正反対です。社会的な公正さはそれほど大きくないですが、全体としての習熟度は非常に低いです。



そして日本は、この2点両方を上手く運営している国です。これは、私たちの研究から得られた非常に重要な結果です。日本はこれを達成している数少ない国のひとつですが、唯一の国ではありません。実際、そうした国々の多くがグリーンエリア〔10位以内〕に到達しているのです。これは、PISAの結論のなかで、政策的に最も重要な結論の1つです。なぜなら、多くの人がどちらかを公平性と質のどちらかを選ばなければならないと考えているけれども、実際には、比較的高いレベルでありながら、この二つの目標を達成でき

る国が、いくつもあるということが分かるからです。質の高さと公平性のいずれも犠牲にしない教育政策なのです。このチャートを良く覚えておいてください。後ほど、これにもとづいて、教育システムについて検討していきます。

金子：今までのスコアの分析を組み合わせると、こういう図ができます。縦軸は成績の絶対水準自体、上の方にいくほど平均点が高い。社会的な背景が成績にどのような影響を及ぼしているのか、影響力が高い国を左、低い国を右、逆にいえば社会的な効率性の低い国が左、高いほうが右のほうというような横軸にしてそれで区切ってみると右上の方は平均点が高いし、全体的な効率性が高い国が出てくるわけです。例えばフィンランド、日本、韓国がそれです。左の上のほうは、平均点は高いけれども、効率性が低い、これはイギリスみたいな国で、右の下のほうはイタリアのように比較的平均点は低いけれども、効率が高いという国というふうに、いくつか分類できるわけです。

3.4 どうやって到達するか

シュライシャー氏：これらの結果に対して、どのような分析的な回答があるでしょうか。まずは、フレームワークをご説明しましょう。その後に質問もお受けいたしますししょう。

生きるための知識と技能			
	アウトプットと 成果 学習のインパクト	政策・施策 成果を規定	前提条件 政策を規定したり障 害となる
国又はシステム	教育の成果	システム全体の組織機構、諸資源、政策	国の教育、社会経済状況
学校や他の教育機関	学校・教育機関の成果とパフォーマンス	学校の学習環境	地域や学校の特性
教授学習場面	教授学習の質	教授学習の実践及び学級の雰囲気	生徒の学習環境や教師の勤務環境
学習者	知識と技能の状況と分布	生徒の態度・意欲・行動	学習者の背景

OECD各国での生徒の習熟度を検討する場合、つまり、国別の習熟度を検討しようとするなら、機関別の習熟度にかかわる問題があります。各機関、つまり、

学校はそれぞれの国でどのように機能するのか。学校による違いは大きいのか小さいのかなど。教室のレベルについての疑問もあります。教室内では何が起きているのか。教室レベルでの質の問題です。さらに個人レベルの問題もある。各国の習熟度を判断したいのであれば、教育システムを、同時に4つの異なったレベルにわけて調べる必要があります。

そして、3つの主要な問いがあります。まず初めに、全体的な結果です。4つのレベルで、それぞれどのような結果が得られたのか。第2の問いは、何を変えることができるのかについてです。それらの結果に影響を与えうる政策とは何なのか。そして、同時に重要なのが、それらの政策のコンテキストとなっている制約は何か。つまり、変えることができないものは何かということです。これは、これらの結果を解釈しようとするときに、心にとどめておかななくてはならないことです。例えば、日本は非常に上手く教育政策を行っているとと言えるけれども、他方で、「日本はお金持ちな国であり、何世紀にもわたる教育の伝統がある。にもかかわらず、その結果を、このような伝統を持たない他国と比較している。だから、それは公正な比較ではない」と論じることでもできるのです。このような部分を常に意識しておかねばなりません。だから、最後の問い「変えることができない前提条件についての問い」が非常に重要なのです。結果を見る際にはしばしば忘れてしまいがちなことですが、それはきわめて重要な問いであると思います。

では、この表に戻しましょう。各国の全体的な習熟度やその国のランク順位に新聞は大変な興味を示しますが、これらからは大したことは分かりません。より重要なのは、機関の機能です。異なるタイプの学校や機関がどのように機能しているのでしょうか。日本ではまだ実施されていませんが、多くの国でははるかに多い最高50000人にもものぼる生徒をサンプルとして、学校のタイプや地理的なエリアによる違いを明らかにしています。機関レベルの習熟度は非常に重要な問題ですが、あまりよく研究できていない分野のひとつです。また、私たちは、授業の質の評価、すなわち授業の比較をどうすればよいのか、実際よく分かっていません。また、生徒たちの背景状況を同等とみなしたとすると、そんなに大きな違いはあるのか、異なる授業で達成される結果の質はそんなに違うのかという疑問もあります。これは、いい問いですが、私たちは、まだ、その

問いに答えることはできません。そして、最後はもちろん、個人レベルでの知識や技能の分布ですが、これは簡単ですね。テストの点数そのままです。

政策といいますか、成果を規定するものとしてはそれぞれのレベルであります。個人レベルでは、構えやモチベーション、関与などが挙げられます。しかし、おそらくこれらは、政策的な介入対象ではありますが、非常に重要な教育の成果でもあるのです。学習戦略も、こちら〔結果の規定要因〕に分類することもできますし、そちら〔教育の結果〕に分類することもできます。PISAでは実際、そのようにしています。私たちは過去の教訓から、学習戦略を、結果としても、また結果の規定要因としても捉えました。正しい学習方法というものは、とても重要な教育の成果だからです。日本では、多くの生徒が学習戦略においてあまり良いスコアをだしていませんね。

そして、前提条件としては、これは簡単ですが、各国のマクロ経済的な特徴があります。それから、学校や学級の背景、そして先ほど言及した生徒の社会的背景です。これがPISAにおける私たちの分析枠組みです。ここまでで何かご質問はありますでしょうか？
金子：PISAの考える分析の枠組をここに整理してみると、縦のほうですね、国やシステム、学校、あるいは教室で何をやっているのか、個々の学習者がどのような問題があるのかという問題がいくつかレイヤーとか、層がある。それから、表等を見ると、学習がどのような成果を上げているのか、それはどのような要因によって規定されているのか、それよりもその前にどのような前提条件が各国にあるのか、といったマトリックスができるわけです。その中でそのマトリックスを考えてみると、このような問題がある。その中でPISAはこれまで分析がよくできているところもあるし、まだそれが課題になっているところもあるということです。

3.5 教育への投資が及ぼす影響

シュライシャー氏：次に進む前に、なにかご質問はありませんか？

人々がまず興味を持つもののうちの一つに、お金と結果の関係があります。これはすでに研究されていることで、実際には興味深いことに、この二つの間にはまったく相関がないとされてきました。しかし、こ

では0.48の相関を示しています。統計は30カ国のみでしか取られていませんが、統計的な関連が生じつつあります。低学年から15歳までを評価した結果です。

日本は、投資金額から予想される以上に、ずっと良い結果を出しています。しかし、期待を下回る結果しか出していないイタリアのような国もあります。面白い問題ですが、あまり信頼できる結果ではありません。
金子：一つの問題は、学校で1人当たりどれくらいお金をかけているのかということで、平均の各国別の成績と相関させています。意外にあるわけですが、はずれている国もあります。

シュライシャー氏：日本では、いつもここで同じ質問をされますが、投資金額の合計において、私的な教育支出が、メインであったというのは事実です。

3.6 学校の自由度が及ぼす影響

シュライシャー氏：直接結果についてお話したいと思います。そう、これがお見せしたかったスライドですね。この図を覚えていらっしゃいますか。教育制度の質と公平性にかんする図です。PISAを通じて分かったことの一つに、結果志向の教育が、結果に大きな影響を与えるということがあります。第一に、結果志向というのは、どの程度生徒たちに向上を期待するかとか、成績に応じた学生番号などとかかわっています。これは、達成するために圧力をかけるということではなく、いい成績を収めることを奨励したり、個別に支援するなどといったことです。これは個人レベルで行われるものです。学校レベルでは、ある種のインセンティブが関わってきます。よい結果を出せるように学校に必要なものを供給することは、学生がよい結果を出せるようにすることなのです。

すると、このような問いが浮上します。現在、大きな政策的議論を呼んでいるものですが、それは、学校の自由度を高めることで、よい結果を出せるかという問いです。すなわち、各校の学習結果の管理、独自の学習環境の設定、教師を雇うための資源配分、そして、教師の解雇など、こうした事柄にかんする自由です。細かく分析してみると、結果はとても印象的なものでした。もちろん、直接的な因果関係など決して想定できません。しかし、これを見てください。

結果の良くない国のシステムでは、意志決定構造が非常にリジッドである傾向があり、学校にはとても限

シュライシャー氏：私たちはOECDの仕事で、多くの時間を費やして、すなわち教育システムや構造管理などを支える特性、いわば駆動力を明らかにしてきました。そしてこのような図を描いたわけです。興味深いことに、このシステムは「フランス?」、最も過剰に規制された構造となっているにもかかわらず、学校間分散が全分散の3分の2を占めています。資金をどれだけ公平に分配しようとしても、結果に違いが生じてしまうということです。

3.7 トラッキング・システムが及ぼす影響

シュライシャー氏：このもう一つの図も非常に面白いものです。

習熟度の分散にかかわる、2つの基本的な取り組みがあります。1つはドイツ、オーストリア、ハンガリー、チェコ共和国といった国々でみられるのですが、そこでは、種別化された試験を実施する学校環境を作ることで、もっとも大きな分散を生み出しています。基本的には、生徒をトラックやストリームに分類するのです。これがまず一つ目の方法です。もう一つの方法の典型はフィンランドやスウェーデンです。特別な学校というものはなく、1つのタイプの学校しかありません。そこでは、生徒個々の違いや、ニーズ、関心、そして能力などに注意を向けるのが、教師の仕事です。これらは2つの極端な例で、他の国はこれら2つの間に位置しています。では結果を見てみましょう。

金子：もう一つ非常に面白いのは、学校のタイプによってもアチーブメントが違うということです。一つは

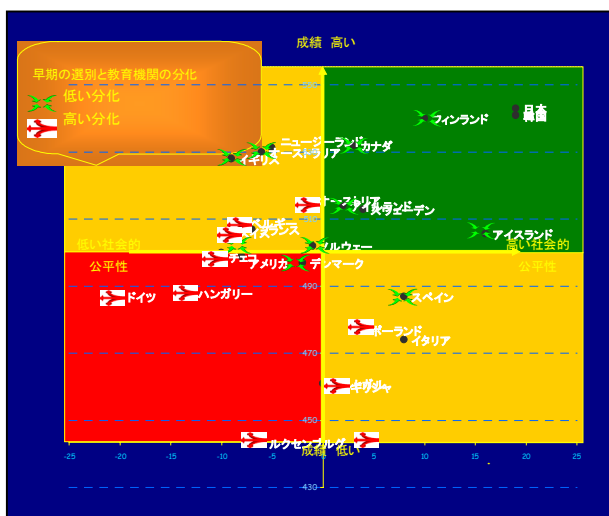
ドイツとかハンガリーのように、かなり早いレベルで生徒を進学トラックや、別のコースに分けてしまう、もう一方のほうではフィンランドのように、学校の全てを分けずに、教師が個々の生徒に対して、より多くの注意を払うという形のシステムもあります。そしてその両端の間にいろんな国が分布しているわけです。

シュライシャー氏：では、結果をご覧くださいませ。この結果に驚くのは私だけではないはずです。私が驚いたのは、これだけです。PISAで明らかにした結果のほとんどは、これまでの研究からおよその見当はついていましたが、この結果には、とても驚かされました。私たちが「統合度が高い」と呼んでいる教育システムに、印をつけてあります。そこでは、基本的には、生徒の個人差に対処するのは教師の仕事であって、生徒をグループ分けする仕組みは特にありません。赤いのは、分化度の高い教育システムです。なお、ここで日本が分類されておられません。公式的な情報に基づくかぎりでは、日本は統合度の高いシステムとして緑で表示されるのですが、実際には非公式にであれ学校内で多くのトラッキングが実施されていると聞いております。そして、その量を数値で表せないため、日本は、この図には入れませんでした。

さて、しつこいようですが、非常に興味深い結果です。最も成功している教育システムは、生徒の能力のばらつきに対応できる教師を抱えています。そして、こちらのように、システムの構造によってばらつきに対応しようとしている国では、目覚ましい結果を出していません。次で最後のスライドとなりますが、この理由の一端をご説明いたします。

金子：生徒をどの段階で、非常に早い段階で分類して、分解して、トラックに分けてしまうのか、そうではないかということで、先ほどの図に分類を分けてみたのがこの図です。比較的分化が遅い国が右上、分化が早い国が左下にあります。日本の場合にはトラッキングがいつ起こっているのか、どのような形で起こっているのかということに議論があるので、あえて入れませんでした。かなり明確に二つのタイプの間に差ができています。これまでの結果は大体今までの学歴調査からも予測されていたことですが、これは非常にある意味では驚きというか、この調査で初めて出てきた面白い結果です。

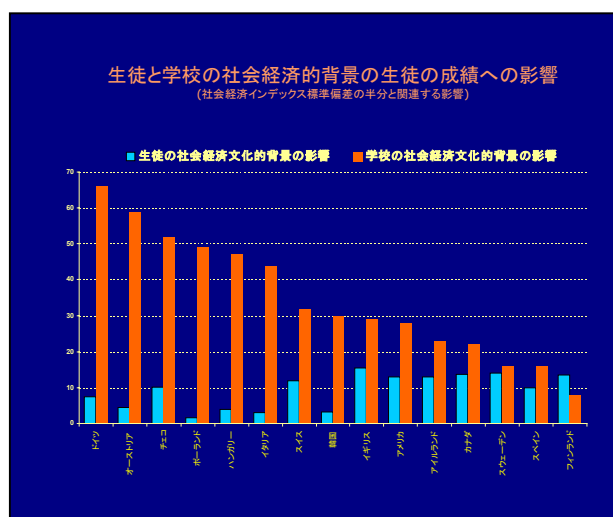
シュライシャー氏：この理由は何なのでしょう。どうやって、これを説明すればいいのでしょうか。生徒



を能力別に分けるのは、個々に適した環境を生み出すのですから、理論的には好ましいはずですが、私もこのように直感するわけですが、皆さんはすでに結果が正反対であることがお分かりでしょう。では、どのように分析すれば結果に説明がつくのでしょうか。

生徒の習熟度に影響を与える社会背景について二つの側面から分析するのです。1つには、生徒の結果と家族背景には直接の因果関係があります。例えば、貧しい家庭ほど、子どもにいい本を買い与えたり、勉強機を与えたりしません。これが生徒の習熟度への直接的な影響です。他方で、間接的な影響もあります。例えば、貧しい家庭出身の生徒ほど、出来がよくないクラスや学校に入るといった具合です。PISAのような研究では、これらの要素を分解することができるのです。

興味深い結果が出ております。青いバーは家族背景との直接的な関係を、赤いバーは間接的な関係を表しています。これで、トラッキングが実施され、早い時期に選抜するシステムが赤い方に偏る理由が分かります。実際には、能力にもとづいて生徒を分けているのではないということです。能力で分けようとしていますが、実際に行っているのは、ほとんど社会的な選別です。人的資本理論に基づいているわけですが、効果はありません。



トラッキングがひとつのインデックスとなっていることは確かです。トラッキングを実施する教育システムがうまくいかない理由を説明することはできませんが、そうした教育システムがみなチャートのこちら側

に分類されるのですから、トラッキングを実施しているかどうか、格差を生み出すシステムかどうかを識別する強力なインデックスとなっているわけです。

金子：なぜ先ほど言ったようなことが起こるのかということです。それは考えてみれば非常に不思議なもので、比較的早い時期に生徒を分けたほうが、それぞれに合った教育が生徒にできるはずではないかというふうに考えられるわけですが、なぜ実際にはそう起こらないのかと考えていきます。するとこのデータから、生徒が直接自分の社会背景から受ける影響と、生徒が社会背景から学校を選択してその学校の影響によって受ける影響がある。その二つに分けて考えると、学校を経由した社会、経済、文化的背景の影響が非常に大きい。これはドイツやそれらの国で生徒間の格差が大きな原因になっています。

4. PISAの将来像

シュライシャー氏：すみません、持ち時間を少し超えてしまいましたね。もしよろしければ、このプロジェクトがこれからどのように発展していくかを簡単にお話したいのですが。

金子：どうぞ

4.1 IT基盤の導入

シュライシャー氏：まず、これは私たちの次の調査サイクルです、2003年調査は現在進行中です。この中で、新しい要素は問題解決スキルです。2006年には把握するリテラシーの範囲を、情報コミュニケーション技術（ICT）リテラシーにまで広げたいと考えています。その理由は、第一に、このスキルがとても重要だと考えるからです。しかし、まだ明確になっているわけではありませんが、これ以外にも重要だと考えるものがあります。

PISAにおける全ての評価を、情報技術基盤 information technology based platformへ移行させたいのですが、それは、効率よく評価が行えるようになるという理由からだけではありません。現時点では、日本の学生でもメキシコの学生でも、一種類の試験で、各生徒のさまざまな能力をカバーしなければならないので、とても効率が悪い。インタラクティブなプロセ

スを導入すれば、もっとよくなるでしょう。そうすれば、試験項目を選択することができるようになり、生徒が試験を受ける際に、正解すれば、より難しい問題を与え、逆に、間違えた回答をすると、次の問題は少し易しいものを与えることができるようになります。このように適応性のある技術を用いることで、生徒の時間をもっと有効に使うことができます。PISAの試験をメキシコの生徒にやらせると、彼らはとてもいらつきます。彼らにとっては、非常にづらい体験なのです。ほとんどの生徒が、どんな問題でも間違えるのです。同じ試験を日本で行うと、生徒たちは多くの問題で正しく答え、自分の能力が高いと考えるかもしれません。適応性のある環境を整えるという意味でも、新しい技術を取り入れるのはとても重要なのです。

また、よりインタラクティブな方法で試験を行いたいと考えるのには、もう1つ理由があります。例えば問題解決能力を評価するとします。この場合、既成の試験を用いてはいけません。出題は、生徒の回答に応じて、展開させていくべきだと思います。試みにこの形式で実施してみたところ、とても面白い結果が得られました。この試験の課題は、生徒にスペースシャトルのエンジニアになってもらうというものです。生徒はある惑星から別の惑星へ行かなければならず、その間に色々なハプニングが起きるのです。このため、ダイナミックな問題解決能力が必要とされます。急に状況が変わり、意思決定を行うと、また違う状況に変化する。これによって、いままでは評価されていなかった新しい能力を見ることができます。読解、数学、科学といった伝統的な能力は、それぞれが強く関連しています。つまり、1つの科目ができる生徒は、他の科目でもいい成績を収めるという傾向があるのです。しかし、ダイナミックな問題解決能力は、全く新しい種類の能力です。だからこそ、ITを導入したいのです。また長期的には通常の試験を行う基盤としても利用していきたいと考えています。

金子：今後何をやるかということですが、2003年以降こういった課題があります。これと同時に試験の方法自体も、もう少し違う方法を考えていきたい。それは基本的にはインタラクティブというか、試験で何か答えたならば、そこからまた新しい問題が提示されて、それに答えていって新しく問題が提示されるというような形のテストです。では、なぜ重要かというと、一つにはやはり生徒の学力レベルによってどんな

学力レベルでも、同じ試験をやっていると非常に大きな問題が生じます。例えばメキシコの生徒と日本の生徒と同じ問題でやると、やはりメキシコの生徒がなかなか答えられなくて、というような問題ができる。それから、動的なというか、ダイナミックなコンテキストで問題を解決するような能力を測っていくテストも非常に重要だということです。それからアカデミックな数学とか科学というのは、大体どのスコアを見ても非常に相関が高いのですが、ダイナミックなコンセプトでどのように問題を解決するかというような能力については、むしろそういったことは分からない、そういったところを測っていくのが重要だと思います。

4.2 対人スキルの評価

シュライシャー氏：これはまだ成功するかどうかは分かりませんが、2009年には対人スキルの要素も組み込みたいと考えています。2006年に開始する計画は2009年には実施されると思うのですが、今のところまだ、どうなるか分かりません。さまざまな課題やトレードオフに直面するでしょうから。課題としてはまず、把握できるスキル領域の幅を広げていきたい。ただし、これは非常に重要な目標ですが、すべての努力をここにつぎ込むわけにはいきません。

最初に説明したとおり、この研究は政策に多大な影響を持っているために、私たちは、教育システムの成否を適切な基準によって評価したいのです。驚くべきことに、このような研究はOECDにはありませんでした。OECD全体を見渡して、教育に関する研究だけでなく、経済に関する研究を含めても。PISAは政策に本当に大きな影響を及ぼしました。日本はいい結果でしたので、それほど影響はなかったと思います。しかし、結果が悪かった国では大きな影響を受けています。私の国では、結果が発表された3週間後に教育の予算が18%増えました。教育システムにこれだけ大きな影響を与えるため、責任を持って評価をするならば、より広範なスキル領域を把握しなければなりません。さもなければ、リスクを冒すことになります。とはいえ、これも全体の中では多くの目標のうちの1つにすぎません。

さらに、対象母集団を検討する必要があります。ここで重要になるのは、生徒の学年があがったとき、私たちも上の学年を対象とすべきかということです。違

う学年の生徒も評価することで、縦断的研究を行う必要があるかもしれません。とても重要な研究目的がいくつもあるので、全力でスキル領域の問題にかかわることはできないのです。誰を評価するのか、小学生も対象にすべきなのかなども考えていかなければならないのですから。

4.3 文化的多様性への配慮

3つめは難しい問題です。

国際比較は、各国の教育システムの長短について知ることができるのですから、大変面白い。ただし、各国の文化的な多様性にはもっと注目すべきだと思います。これについては、すでに取り組みを開始しています。例えば、PISAで出された課題はすべて各国によって評価されています。各国の難易度の範囲をカバーしているかどうか、課題はカリキュラムの面からは適切かなどが検討されます。主観的ではありますが、すべての課題がこのように評価されます。次にこれらを客観的にテストします。サイコメトリックな特徴に注目するのです。試験項目は国にかかわらず同じように機能するのか、それとも異なるのか。相対的な難易度は、すべての国で共通しているか、文化によらず共通しているか。これは非常に重要な検討で、こうした検討を踏まえて、それぞれの課題を採用するかどうかを決定するのです。そして、これは国と国だけの問題ではありません。それ以上に一国内の差異にもかかわっています。さまざまな集団、さまざまな教育プログラムを考慮に入れたとき、このテストは、本当に1つのものを測定していることになるのでしょうか？ それともさまざまなものが測定されて、その平均値だけが把握できている状態なののでしょうか？ これに関しては、これまでもいろいろな形で取り組んできましたが、もっと注目していい問題だと思います。また、問題解決能力、コミュニケーション能力、対人能力などといった、評価するスキル領域を広げるにあたって、より重要となってくる問題だと思います。領域を広げれば広げるほど、この問題には、さらに注目していかなければなりません。

金子：今後の問題点ですけれども、一つは、より広い分野での能力というのに調査の対象を広げていく。2009年には対人能力みたいなものも対象に入れることも考えていますけれども、重要なのは、広げるだけで

はなくて、それと同時に概念もしっかり考えていかなければいけない。それはなぜかという、このテストは多分OECDのいろんなプロジェクトの中でも非常に政策的な影響力が大きいものだからです。例えばドイツの場合にはこの結果が発表された後に、連邦政府の教育支出が18%増えたそうです。そういう点で非常に直接的な影響が大きいので、これからは慎重に基礎的な概念を考えていかなければいけません。

それからもう一つは調査対象ですけれども、今15歳だけでやっているわけですが、先ほど言ったように追跡調査、あるいは小学校などで調査をすることも重要だろう。

もう一つ。国際比較は、非常に意味があるわけですが、国によって調査項目が捉えているものが非常に違うのではないかと。そういった意味で調査項目が国によって、文化的なコンテキストでどのように違う捉えられ方がなされているのか。これまでもいろいろと分析していますが、さらに、分析していくことが必要だろう。

4.4 政策的側面と学術的側面

シュライシャー氏：これで終わりたいと思いますが、私たちがどのようにPISAを運営しているかについて、少しお話ししましょう。私たちは、政策とテーマを分離するのに大変苦労しました。まず、プロジェクト全体を管理するための、参加国からなる委員会を設けました。そして、この委員会が、PISAで何が重視されるべきかを戦略的に判断するわけですが、彼らが期待する研究内容も国によって違うでしょう。

また、別の問題ですが、私たちは科学的な基盤に依拠しています。世界各国の先進的な機関と共同でテスト開発を行ったり、OECDのために分析方法を開発したり、政策の基礎となる科学的な概念を開発したりしています。これにはさまざまな専門家が携わっています。数学、科学、読解、問題解決や、精神測定などさまざまな分野の専門家が関わっています。したがって、OECDと実際に仕事を行う機関および個人がネットワークを構成しているのです。私の部署は政策と技術を結びつける役割を担っています。技術面は科学的な面で起こっていることに影響を及ぼします。プロジェクト全体が方法論への関心のみにによって支配されているということはありません。方法論は、政策的課

題に回答を与える際に用いられる補助具だからです。しかし、逆に、方法論への関心が希薄であれば、これもまた危険です。現在では、とても手の込んだ方法論を用いることも可能ですので、節度をもたねばなりません。

プロジェクトは、以上のように動いております。では、この辺で終わりにしたいと思います。議論やご質問があればどうぞ。

金子：最後に申し上げておきたいのは、このプログラムは政策的な側面と、それから学術的な側面の二つがあるということです。政策的な側面ではOECDに理事会があって、それによってプログラムの方針が決められます。もう一方の学術的な側面では、世界各国からさまざまなエキスパートが集まってもらって議論しています。ただ、学術的な側面だけが強調されてしまうとまた問題があるので、両方の側面について注意しながら決めていくことが非常に重要だと思います。

5 質疑応答

金子：質問、ディスカッションをしたいと思います。質問、ディスカッション、いずれも英語が好きな方は英語でやって頂いて、他の人のために多少私が要約したいと思います。あるいは日本語でやりたいかたは日本語でやって頂いて結構です。ご自由にどうぞ。

質問者A：PISAの調査で測定されている能力というのがかなり抽象的なレベルで、そこで、15歳で測定されているのだけど、それがどの程度学校で作られるのか。それとも他のいろいろな社会状況を含めて、あるいはマスメディアも含めて、どの程度学校外の要因がかかわっているのか。そこら辺のコントロールはどのようにされているのか。

金子：私たちはPISAプログラムで測定しているものの定義をうかがいましたが、測定されているのは、抽象度の高い概念のようです。そこで質問ですが、これらの能力はどの程度、学校とそれ以外の環境、すなわち、メディアや家族などによって形成されるのかということです。これらの分離は行っているのですか？

シュライシャー氏：2つを分けるための評価は行っていないので、この質問にははっきりと答えることはできません。数学的リテラシーと比べると、おそらく読解リテラシーの分野の方が、学校以外の影響が大きいと

思います。

しかし、私は学校が結果に与える影響は大きいと考えています。まず、学校間で大きな分散があるということです。結果が学校外の影響しか受けないのであれば、学校間のこれほどの分散は見られないはずです。しかし、現実には学校間で大きな差が現れている。また先ほど説明したとおり、学校間の差のうち70%は、学校の背景的な特徴によって説明ができます。したがって、結果と教育システムの構造には強い関係があるのです。これらを考慮すると、学校が大きな影響力をもっていると言えます。

金子：学校が作る能力、学校以外のものと二つに分けるのは非常に難しいと思うのですが、ただ一つこのデータからいえるのは、学校間の差はかなりある。非常に大きい。そのバリエーションの70%は、学校によって説明されるわけで、それだけ大きい。それからまた国によってもかなり差がある。それは要するに、ただ単に他の環境的な影響が非常に大きいのであれば、学校による差はそれほど大きく多分出てこないだろう。学校によってかなり影響が大きいのではないかと思う、ということです。

シュライシャー氏：もう1つ、伝統的な科目ではなく、こうしたスキルの習熟度を測定する理由について付け加えておきます。理由の1つめは当たり前とも思えるものなのですが、私たちが実施したほかのリテラシー調査によると、これらの概念的能力、つまり、知識を生み出すための情報処理能力が、学習能力を大きく決定づけるのです。また、これらは個々人の将来生活を大きく左右します。他方で、従来の科目別能力は、国際調査によれば、大人になってからはあまり役立たないことが明らかとなっています。だから、私たちは、これらのスキルを重視するのです。

金子：もう一つ、これに関連して申し上げておきたいのは、なぜ我々がもっと伝統的な学校で教える学科の学力に重点をかけてテストをしなかったかということです。それはアダルトリテラシーの調査を別にやったわけですが、それはやはり学科の能力よりも、もう少し抽象的なコミュニケーション能力、そういったものにむしろ非常に大きく影響している。だからそちらの方が重要ではないかというふうに我々は思ったわけです。

質問者B：興味深い発表でした。これらの変数は、どのように評価されたのでしょうか。クラス内の状態を

評価するためにビデオを使用したとおっしゃっていましたが、生徒の評価はどのように行われているのでしょうか。実際の試験について、よく知らないのでお尋ねしたいのですが、たとえば、私が多肢選択型の試験をすべてできたからといって、それは私のリテラシー能力を測定しているとは思えない。また、他の方法、例えばエッセイ式のテストを行ってスコアを出すという場合には、答えに対する適切なスコアを出すのは大変難しいと思うのです。このことについて少し教えて欲しいのですが。

金子：日本語でちょっと今の質問して頂けますか。

質問者B：リテラシーという言葉にはすごく興味深いものがたくさん含まれていると思うんですけども、それをどうやって測定するのかというのがとても難しいと思います。例えば、クラスの条件とかならば、ビデオテープを使ってやったりもしているとおっしゃっていたんですが、子どもの場合はどうなのか、ビデオを使っているのか、そういうことが聞きたいと思いました。もしも多肢選択だけだったならばそれは正しく測れないだろうと思うし、もしエッセイ式のようにテストで測るとしても、それをスコアリングする時に大きな問題があるような気がします。その点についてどう考えているか聞きたいんですけども。

シュライシャー氏：そうですね、私たちとしては、できるだけ多肢選択型の問題を採用したかった。しかし課題を作った専門家によれば、多肢選択型では私たちが評価したい能力は測れないそうです。このため、自由回答の問題を多く採用しなければならませんでした。物事を評価、分析、比較するような問題の場合はとくに自由回答形式を採用しています。

また、ご質問の通り、これらを採点するのは大変難しい。まず、詳細な採点基準を設定しました。次に、試験的に、その基準にそって採点を行うために、生徒の回答をたくさん集めます。ここまでは比較的簡単です。大変なのは、実際の採点です。信頼できるスコアを得るために、自由回答の問題については、4人の採点者によって採点しなければならなかったのです。そして、採点者間で信頼性を評価しあいました。しかし、各国は採点者の文化的態度といった要素がかかわるのではないかと懸念していました。例えば、採点の荒さは国によって異なるかもしれません。そこで、一部の項目については、私たちは生徒の回答を各国用に翻訳した上でさまざまな国に送り、採点しなおすよう依頼

しました。三角測量のようなやり方です。このようにして、自由回答の採点について、採点者や国の間で信頼性を高めることができました。非常に大変な作業でしたが、有意義でした。何が上手くいき、何が失敗するかもわかってきました。基本的に、採点者間の信頼性はとても高いことは分かっていました。採点手順がとても複雑なため、すべての採点者には研修を受けさせたからです。すべての国の採点者が非常に特殊で詳細にわたる手順についての研修を受けます。したがって、不明確なところは、ありません。これは大変上手くいきました。国の間での信頼性については、調整を要する問題がありましたが、調整することができました。採点者は常に荒いかに細かいかのいずれかなので、調整するのは比較的簡単だったのです。

採点の秘訣は、とても緻密な手順と、しっかりとした研修です。専門家によれば、多肢選択型の問題を多く使ってはいけません。50%くらいはいいかもしれませんが、残りは自由回答型でなくてはいけません。

金子：試験の方法についてですが、選択式をやったほうが楽というか、いいのかもしれませんが、専門家はあまり選択式が多いといけないということで、記入式の試験、記入式の部分をかなり多くしたわけですね。そうするともちろん採点上の問題が出てくるので、そのために採点のガイドみたいなものも作って、かなり周到に採点の注意をすると同時に、一つの国で同じ試験を4人で採点することにしました。もう一つは、国によって採点に厳しい国とそうでない国が出てくるようなので、これは一定の質問について一つの国でやった回答をまた他の国に送って、そこで採点してもらうこともやりました。採点者による違いは少なく、国による違いがむしろ大きな問題であったということです。

シュライシャー氏：この本には採点の手順がかかれています。この本はここに置いていきますね。採点者間の信頼性の詳細な説明と分析がかかれています。

金子：これは、最新の報告書だそうなんですけれども、どうやって採点したのか、そのテクニカルレポートが出ています。

質問者C：日本はとても平等だというんですけども、私が調査をやるといつもバックグラウンドを聞くのが非常に難しく、リジェクトされることが結構あるんですが、今回のサンプルでそういうことがあったのかどうか。

シュライシャー氏：これはいい質問ですね。正直に言

いますと、日本版にはそれらの質問はなかったはずなので、生徒がどう回答したのかは分かりません。決定的に重要な質問項目は、親の職業ですが、生徒はこの質問に答えることを許されなかったのです。親の職業は重要な変数だと思うので、大変残念です。そこで私たちは、社会背景に関係する他の変数を用いて、日本の変数を代替しました。非常に強い相関関係がみられましたので、文化的活動、社会活動、家族の裕福さなどの質問項目は、とても信頼性のあるものだと思います。それらは、社会的背景にかんする質問の、非常にすぐれた代替項目ではありますが、従来のやり方で評価を行えないのはとても残念です。なぜなら、両親の職業は社会的背景を予測するのに最適な尺度だからです。文化的な行動や、社会的活動よりずっといい。

日本の結果は、誤差が他の国と比べて3倍です。回帰分析の予測が完全に安定しているわけではないため、社会的背景変数の誤差は3倍となっているのです。しかし、私が思うに、どんな方法を用いたとしても、日本は公平な社会として位置づけられるでしょう。韓国は日本によく似た国であることが知られておりますが、韓国では親の職業に関するデータはほぼ完全にそろっているのです、信頼性のある結果だと思います。

金子：日本の場合には、両親の職業を聞くことができなかったのも、他の質問を聞いてそれでもって両親の職業代わりにしたわけですが、相関係数はかなり高かったのも、代わりの変数としては一応大丈夫なのではないか。それから韓国もほとんど似ています。韓国は両親の職業を聞いているんですが、そこほとんど同じ様な結果になったわけです。

質問者D：ちょっと関連で、そのソーシャルバックグラウンドを横軸にとり、縦軸に生徒スコアという相関がありました。そこで回帰直線がフラットだとそれは公平でいい社会だと、回帰直線が立っているとソーシャルディスパリティーと言っておりましたが、そうすると回帰直線がフラットなのが、理想的な国なのか。その中で、今言われたのは、ツールアビリティー、本当の能力のみが生徒スコアに反映されるべきだといわれたんだけど、ソーシャルバックグラウンドを・・・[聞き取り不能]・・・が高いということは、何らかの教育的な価値を持っているわけで、そこで能力が高められたものはツールアビリティーではないのか、何を測ろうとしているのかということの質問があるのですが。

金子：社会的背景と試験のスコアを関連づけるこの図に関しての質問だと思います。この線がフラットであればあるほどよく、平等な社会であるとおっしゃっていましたが、ここに入っている人が、社会的背景が高いということはないのですか。したがって、この中に何かがある。社会的背景が高い生徒の能力や社会的スキルが高くなるということはないのでしょうか。社会的背景が高いというのは、何らかの利点を含意しています。したがって、問われているのは、測定されているのは何かということなのです。それと社会的価値とはどのような関係があるのでしょうか。また、測定されている能力を構成しているのは何なのでしょう？

これでいいですか？

質問者D：真の能力さえもが社会的背景の影響を受けているとき、「真の能力」という言葉で何を意味しているのですか？

シュライシャー氏：よい質問ですね。教育システムは生徒のポテンシャルを最大限に引き出してやるべきです。そして、生徒の生育環境などとは関係なく、そうあるべきでしょう。生徒が自分の最大限のポテンシャルをいかせることを保証するのが、よい教育システムや社会システムであると私は思います。あなたは、立場を示せとおっしゃるのですね。

さて、社会的背景そのものは、この関係の中で重要な変数ですが、恐らくそれは1つの要素でしかありません。だからこそ、最も公正な教育システムにおいても、成績の回帰直線の傾きはフラットにはならず、国によって異なっております。傾きを比較的フラットに抑えることに成功しているフィンランドなどは、達成可能な政策目標を示しているように思えます。しかし、私はそうは思わないのです。他の選択も可能ですから。いずれにしても、ここにある結果によれば、教育は、社会的背景と習熟度との関係に大きく影響を及ぼしているのです。ただし、どの程度達成できるのか、どの程度フラットにできるのかはまだわかりません。

質問者D：多くの意味を持っていると思っていましたが、しかし思っていたよりもはるかに複雑ですね。

金子：答えのほうは、多分、政策的に関係をなくしたほうが望ましいという立場から見ているということだと思います。それは多分見方によって違うので、荻谷先生の異論はその回帰直線の傾き方とかいろんなことを意味しているんだろうということだと思います。

質問者E：非常に斬新的な試みだと思います。そのう

ちの一つがインテグレイテッドなロングタームのプロジェクトで、国際的に、文化や社会的な背景を含めて、ロングタームで見ようとしているということですね。20世紀末から現在にかけての一つの面白い現象というのは、男女間、ジェンダー間のいろいろな教育の機会または発展における差だと思うんですけども、ジェンダーにかんする指標というのはどういうふうなものを使われているのか。

シュライシャー氏：もちろん、習熟度の性差についても注目しました。また、学習への関与や、態度、やる気などのジェンダー差にも注目し、変数を作成するとともに、この変数と習熟度との相互作用についても検討しました。これまでのところ、こうしたことについて分析してきました。

知見としては、読解力は女子の方が大幅にいい。数学では男子の方が少し上で、科学では同じくらいですね。私たちの関心を引きつけたのは、性差が国によって大きく異なるということです。もし、国によって違いがなければ、性差は遺伝的なものと言えるでしょう。女の子は生まれつき読解力があり、男の子は生まれつき数学が得意だと。しかし、ジェンダー間のこれだけの差は、教育とはいいいませんが、社会が大きくかわっていることを示唆しています。また、結果があまりよくなかった国では、性差が大きく、習熟度の低い男子の読解力は言い表すことができないほど悪いのです。最も能力の低いレベル1を見てみると、多くの国で、男子の割合が2倍から3倍になっています。では、数学では反対に、最低水準の女子の割合が男子より多いのかと思えばそんなことはなく、同じ問題がみられます。平均値で男子が女子を上回るのは、飛びぬけてできる男子がいるためで、最低水準を見てみると、やはり男子の方が多くなっているのです。

次に、関与や態度についての結果です。学習戦略やメタ認知において女子は男子を上回っており、大きな性差がみられます。また、これが読解力の結果と強い関係があると推測されます。

私たちが検討している第三の問いは、非構造的な変数と性差との相互関係にかかわっておりますが、まだ十分な研究は行われていません。興味深く、また印象深いことですが、学校が性差を消すことができず、むしろ、性差を媒介している国もあります。この点に関しては日本はあまりよい結果を出していません。日本では性差は比較的大きいといえます。

金子：一般的に女の子のほうが成績がいい、数学では男のほうがいいといったことは一般的にいわれます。ただ、その分析をやっていて一つ気がつくのは、国によってかなり差があるということで、一定の国は性による、ジェンダーによる差が非常に強く出ている国と、それをかなり中和している国がある。日本は比較的まだ差が強く出ている国であるということが違うところである。

それから、パフォーマンスとアティチュードというモチベーションのところをみると、モチベーションのほうでもかなり違いがあって、女の子のほうが比較的モチベーションが高いというか、ストラテジーをもっている。それはリーディングスコアにも影響を与えているということがあるのではないかな。それから、もう一つやろうとしたのは、教授法との関連があるのではないかということですけれども、それは今勉強中ということです。

シュライシャー氏：長期的には、私たちは、二つの点に関心を持っています。ひとつには、こうした性差は時間とともにどのように変化していくのかということです。差は小さくなるのか、それとも、より大きくなるのか。もうひとつは、縦断的調査についてです。習熟度の差異が高校以降の教育further educationや大学、職業の選択にどのように影響するのかを明らかにしたいのです。15歳の男女生徒のこれからの成長を追うことで、スキルの違いが将来、人生における選択にどのような影響を与えるかを見ることができます。パターンの1つは皆さんがよく知っているものです。日本は好例なのですが、数学やコンピューターサイエンスを勉強する女性の割合をみてください。また、生物を勉強する男性の割合をみてください。ここには大きな差がみられます。この差は、学校における能力パターンとどのような関係があるのでしょうか。これは大きな問題です。長期的に見ていけば答えが出るかもしれませんが、時間がかかるでしょう。

金子：長期的な問題として大変面白いのは、ジェンダーによる差が長期的にどういうふうに変わっていくのか、縮まっていくのか、あるいは広がっていくのかという問題。それからジェンダーの差が男女別の進路選択にどういう影響を与えていくのか。例えば、自然化学系の分野は女の子の学生が少ないわけですけども、それは将来どのように変わっていくのかといったことも大変面白い問題だと思う、ということです。

質問者 F：先ほど学校内のちらばりと学校間のちらばりのことについて、学力全体での話がありました。今回の学力は読解と、それから数学と科学とそれぞれあるわけですが、学校間のちらばりというものがその内容によって違うのかどうか、そういう傾向について教えてください。

金子：差を学校内の差と一方では学校間の差に分離しておられました。また、読解力、数学、科学の3つの分野を強調されていました。教科別にみると、この2つの分散のパターンには有意な差はあるのですか？

シュライシャー氏：概して、差はありません。特定の国について、この質問にお答えすることはできませんが、一般的には、国にかかわらず、分散のパターンは同様です。このため、教科別の分析は行わなかったのです。国によっては差があるのかもしれませんが、分散のパターンは、基本的に似ていますね。数学は学校の教え方に影響されやすいので、少し大きな分散がみられます。読解力は学校外の影響をより受けるので、パターンは非常によく似ています。

金子：バリエーション全体では数学が多少大きい傾向にあるんですけども、この二つのイン・スクールとビトウィーン・スクールとの配分というか、分布という観点からは全ての国を全部みるとほとんど同じパターンだが、国によって違うパターンができてくるかもしれません。

質問者 C：2つ目の質問なのですが、日本と韓国で、とても得点が高かった、その二つの国はよく暗記中心の詰め込み教育だと批判されているわけですが、その二つの国がこのPISA調査で、リテラシーの得点が高いその原因は何なのか。つまり詰め込んだって、そういう本当の能力がつくということなのかどうか。

シュライシャー氏：興味深いことに、そうした批判は日本でしか聞いたことがありません。日本人以外からはそのような意見は聞いたことはないのです。

しかし、フィンランドが、学生の学習時間は半分で、詰め込み教育もせずに、日本と同じ結果に達しているのは興味深いことです。違う方法で同じ結果にたどり着いているのです。フィンランドは学外試験を行いませんし、学校の外にインセンティブがありません。違うのは、教師なのです。本当に。学校内のトラッキングは、非公式のものさえありません。教師がクラスのすべての学生1人1人と向き合うことで、フィンランドの学校は、とても高い水準を保つことに成功し

ているのです。しかも、最高水準の習熟度にある生徒の割合が、最も大きい。要するに、こういうことです。日本の教育はうまくいっており、非常に素晴らしい結果を生み出しているのです。世界各国のリテラシーについてのデータがありますが、ここで測定されている能力は、人々にとって、経済的に重要であるものです。日本の教育制度は世の中で重要となっているスキルを授けるという点で大きな成果を挙げています。しかし、問題なく、同じ結果を得ているシステムも他にあるのです。

金子：日本が詰め込みというのは日本に来ると聞きますが、他ではあまり聞きません。日本のテスト結果がいいのは必ずしも偶然ではないというか、やっぱりこれは非常に重要なスキルを作っているということは事実だろうと思います。ただ同じ結果を、例えば、フィンランドが同じように高い結果を出しているわけですが、ここでは別に試験地獄でそれを作っているわけではないし、詰め込みでやっているわけでもない、別の方法もあるということがいえるのではないかということです。

質問者 G：関係していると思いますけれども、教育の内容、それからメジャー、測定されている能力との関係です。今の非常に面白いデータが挙げられたんですけども、その多くは教育のさまざまな条件、例えば社会階層、あるいは学校がどのように組織されているか、そういう条件とここで測定されている学力との相関関係。しかし、先ほど言われたように、その伝統的な学科とは違った内容の能力もここで評価されているわけですから、学校と学力がOECDで測定された能力に関係あるとすると、実際に学校で教えられている内容と、それとここで評価されているものとの関係というのは、果たして考えられているのか。

金子：習熟度を社会背景やスコア等との関連において分析なさっておりました。しかし、問題は、測定しているものが、必ずしも伝統的な学科ではないということです。学校は伝統的な学科の知識を教えようとしているので、ここでは測定されているスキルと教えられている内容には多少のギャップがあるかもしれない。

シュライシャー氏：いいご質問ですね。このことは、問題解決スキルに迫ろうとする場合、より重要な課題となります。たしかに、学校はそれを直接教えるのはありませんが、教えた教科内容から何らかの結果が生じるものです。また、教科内容とスキルの間には不

一致もあるでしょうが、教育制度の評価は、やはり、人々にとって重要な事柄を基準とする必要があると思うのです。そのため、私たちがPISAに取り組みはじめたときから、教えられたことを生徒たちがどれだけ身につけたのかを評価することには、まったく関心がありませんでした。もちろん、こういった方法をとることもできますし、このような研究も過去にはありました。

私たちが興味を持つのは、将来の成功に役立つスキルを、どの程度生徒が習得したかということです。習得の手段は問題ではありません。例えば数学を教えれば、生徒たちの情報処理や問題解決の方法に影響を及ぼすことは間違いないでしょう。しかし、私たちはそれよりも、最終的な結果、つまり問題解決スキルがどれくらいかに興味を持っているのです。哲学的な問題かもしれませんね。

私たちは、当初からカリキュラムには注目していません。もちろん、実際には、あらゆる国のカリキュラムを検討したのですが、理論的なフレームワークに基づいてスキルを定義しています。その上で、私たちは、このテストがカリキュラムに適合させられると言えるのです。しかし、これは、最初に申し上げておくべきでした。カリキュラムについてお話ししなかったもう一つの理由は、国際的な評価をする際カリキュラムに配慮すると、各国にすべてに共通する、最も小さな共通要素に注目する傾向があるからです。「この国ではそれは教えていません。だから、それはテストしません。」ということになるのです。結果的に視野がとて狭くなってしまいます。そのため、私たちは最初から「各国のカリキュラムを忘れよう。白紙からスタートしよう。どの国でも、重要とされる能力はなんだろう？」と言いつづけてきました。

金子：一つの技術的な理由は、カリキュラムで教えていることにだけこだわってしまうと、各国で国際比較をやるときに、どこでも教えているようなことしかテストできなくなって、結局そうすると視野が狭くなってしまいう問題があるということです。しかし、もっと重要なのは、確におっしゃるように、学科としてやっていることと、私たちが測ろうとしている能力との間にはギャップがあるんだけど、それはしかし関係ないことはないだろうと。ただPISAプロジェクトでどちらに関心があるかといえば、やっぱりそれは私たちがいつているところの能力、基礎的な能力であ

ると。そこをやるうというのがPISAのプロジェクトのそもそもの考え方だということだと思います。

シュライシャー氏：これはとても重要なポイントだと思います。なぜなら、PISAテストがいくつかの学校にとって不公平であるという議論も可能だからです。これは、私の国、ドイツでなされた有名な議論なのです。そして、私はこれも正しいと思います。ドイツではナショナル・テストと国際的なテストが実施されましたが、ナショナル・テストでは、PISAの中で評価の対象となっていないけれども、ドイツで教えられている多くのことが評価の対象となりました。そして、人々は、ドイツ人はそれほど出来が悪くないと考えるのです。たしかに、こうした議論も可能です。しかし、これはカリキュラム・テストではないので、生徒たちが習ったことを評価するものではありません。

これは非常に重要なポイントです。なぜなら、公平性にかんする議論は、教育制度の責任にかんするものであって、生徒たちの責任にかんするものではないからです。ある国の教育制度が、各国が重要だとみなすスキルを身につけさせることができない場合、問題があると私たちは認識しますが、それは制度の問題であって、生徒の問題ではありません。生徒たちは、私たちがテストする内容を実際習っていないかもしれません。教育システムの責任なのです。しかし、その責任を生徒に負わせる国もあるのです。

ところで、評価は15歳の学生を対象に行われていますが、国によっては学校に通いはじめるのが7歳であったり、3歳であったりとさまざまです。これを「不公平だ」という人もいるでしょう。国によって4年も長く学校で勉強していることになるのですから。しかし私たちは「それは、国が選んだ政策だ」としかいえません。子どもを3歳で学校に入れば成績が上がるというのなら、他の国も同じようにするでしょう。フィンランドでは学校は7歳から、フランスでは3歳からです。フィンランドではいい結果がでていますが、フランスはフィンランドよりもいいのでしょうか？これが答えです。問題はそんなに簡単ではありません。基本的に、私たちは最終結果を評価しています。そして、政策の有効性、カリキュラム、入学制度あるいはトラッキングなどを変数として重視しています。

金子：この問題に関連して一つの議論は、学校で教えていないことをテストすることはアンフェアじゃないかという言い方もされます。例えばドイツでテストす

ると、ドイツのカリキュラムに合ったテストをすれば非常にいい結果が出るけれども、このテストだとあまりよくない。それはやはり不公平ではないかという意見もあります。

もう一つは、国によっては同じ15歳の時点で試験をしても、それまで受けている学校教育の内容や長さも違うといったこともあるだろう。それは個々の生徒がどれくらい勉強しているのかということに私たちは興味があるわけではなくて、どれくらい学校制度自体が私たちの考える能力を作っているかどうかということに私たちは興味があるし、そういう意味でその政策的な観点からいえば、私たちの方法というのがやはり望ましいのではないかと考えています。

シュライシャー氏：これはとても面白い議論だと思います。また、この他にも多くの議論や質問等があるかと思いますが、皆さんをあまり疲れさせるのも何ですので、この辺にしておきましょう。

金子：今回も日本は4日間か5日間しかいらっしゃらなかったそうで、大変きつい日程なので、今日はここらへんで。

ご発表、どうもありがとうございました。

シュライシャー氏：統計値など、なにか必要なものがある場合は、Eメールでご連絡ください。私の電子メールアドレスはこれです。本当にお気軽にご連絡いただければと思います。

金子：われわれにとって、視野の広さをお持ちであることが非常に印象的でした。単なる大規模なテストなどというシンプルなものではないのですから。私たちCOEの基礎学力に関するプロジェクトが学ぶべきところだと思います。シュライシャーさんとはこれからいろいろとご協力いただくことも…

シュライシャー氏：こちらこそ。私たちの仕事は国々の経験の上に成り立っています。私たちがすべてを發明しているわけではないのです。各国が持つ専門知識は大いに活用していますし、各国で行われている評価とOECDの評価は、お互いに影響を及ぼしあっています。後知恵ですが、以前私がいたIEAとOECDはなぜこんなにも違うのかを説明することができます。OECDは、各国の専門家の意見を集約してきたのであって、専門分野を開発する小さなセンターを設立してきたのではないのです。私の部署では、現在3人の職員がPISAに取り組んでいますが、基本的になにもしていません。専門家も私たちも意図的にそうして

います。そうして、各国の専門家のネットワークを維持しているのです。こうすることで、常に最新の情報を手にすることができるので、私たちはそれ以上の研究はしていません。お役所をつくるのではないのです。これが上手くいっております。

金子：OECD、私たちのCOEも是非勉強させて頂きたいというふうに言いましたけれども、一つはOECDの側もそんなに沢山スタッフがいるわけではなくて、いろんなところとネットワークを作って事業を進めているので、私たちともこれから協力をさせていきたいということです。

本当にありがとうございました。