

理科の活用志向性尺度の作成と信頼性・妥当性の検討 —スーパーサイエンスハイスクール指定校の取り組みの評価を見据えて—

原田 勇 希
秋田大学教育文化学部

草 場 実
高知大学教育学部

齋 藤 恵 介
高知大学大学院教育学専攻

Developing a Science Utilization Orientation Scale to Evaluate Efforts of Super Science High Schools

Yuki HARADA*¹, Minoru KUSABA*², Keisuke SAITO*³

*¹Faculty of Education and Human Studies, Akita University

*²Faculty of Education, Kochi University

*³Graduate school of Integrated Arts and Sciences, Kochi University

In science education, it is important to make students realize the usefulness of science, and Super Science High Schools (SSH) are supposed to foster the attitude of students to utilize scientific knowledge and perspective proactively. It is therefore necessary to evaluate whether the efforts of SSH schools successfully foster expected attitudes. This study aimed to develop a science utilization orientation scale to assess the efforts of SSH schools. The results showed a factor structure as assumption with sufficiently high factor loadings. There were no DIF, ceiling effect, and floor effects, that would impair the validity of the scale, that we found in the results. The difference in the mean value between each group was consistent with the theoretical hypothesis. In conclusion, this scale is considered to be able to measure the science utilization orientation of SSH students and evaluate the efforts of SSH schools.

Key words: science utilization orientation, Super Science High school, SSH, psychological measurement

1. 問題と目的

1. スーパーサイエンスハイスクール指定校における教育評価と理科の活用志向性

我が国の高度な理数教育を牽引する制度であるスーパーサイエンスハイスクール（以下、SSH）事業では、「生徒の科学的能力及び技能並びに科学的思考力、判断力及び表現力を培い、もって、将来国際的に活躍し得る科学技術人材等の育成」を趣旨としている（科学技術振興機構，2019）。また、その事業目的は「科学技術人材の育成の初期段階として、高校生の理系科目への好奇心を育て、科学技術方面への進路選択をサポートすること」と「旧来の学習指導要領の改善に資する試みの実施」の2点に集約される（小林・小野・荒木，2015）。そのため、SSH事業に指定された高校

（以下、SSH指定校）にとっては、理数教科の高い学力（i.e., 知識及び技能、思考力・判断力・表現力等）を身につけさせることのみならず、科学技術やその基礎にあたる教科である理科の学習内容や考え方を、高校生時点から将来にわたり継続的に活用しようとする態度（i.e., 学びに向かう力・人間性等）を育成することも重要な使命である。同時に、SSH指定校は教育課程の改善に資する実証的資料の提出が課されていることから（科学技術振興機構，2019）、教育効果を客観的に評価し、それに基づいた教育改善に努める必要がある。

とりわけ、将来にわたって理科を活用しようとする態度に注目すると、これを正確に評価するためには卒業生を追跡し、理系大学・大学院への進学率や科学技

術関連職への就業状況などを調査する必要がある (e.g., 小林ら, 2015). しかしながら, 卒業生を対象とした長期的な追跡調査には大きなコストがかかる. また追跡調査では教育効果の評価に5~10年単位の時間がかかることから, 即時的または数年単位でのカリキュラム・マネジメント (中央教育審議会, 2015) の確立が難しい. そのため, SSH 指定校では簡便な質問紙法により理科を活用しようとする態度を評価するためのツールが望まれる.

ところで, SSH 指定校が提出した教育課程の改善に資する実証的資料は, 「研究開発実施報告書」として Web 上に一般公開されている (2020年9月現在, 平成14年度から令和元年度まで公開されている; 科学技術振興機構, 2020). 当然ながら, 各 SSH 指定校の教育効果を適切に可視化し, さらに相互に比較可能な実証的証拠とするためには, 信頼性・妥当性が確認された尺度を複数の高校で統一して用いることが不可欠である. しかし「研究開発実施報告書」において, 各 SSH 指定校が示している各指標を概観すると, 各学校が独自に作成した信頼性・妥当性の検討を経していない項目群を用いているケースが多い. そのため, 現状では, 各 SSH 指定校の教育効果が適切に可視化されているとも, それが相互に比較可能なものであるともいえない.

こうした背景を受け, 本研究では SSH 指定校の取り組みによる教育効果を質問紙によって評価することを見据え, 高校生時点から将来にわたって理科を活用しようとする態度 (以下, 理科の活用志向性) の測定尺度の開発を目指す. 尺度の作成にあたり, 以下では, 理科教育と心理学の視点から理科の活用志向性の概念規定を行う.

2. 理科教育に関する調査結果に基づく“理科の活用志向性”の概念規定

SSH 指定校に限らず, 理科の有用性を実感させることは理科教育の重要な教育目標である (e.g., 文部科学省, 2018). しかしながら, 国際学力調査における質問紙調査の結果等を受け, 平成30年告示の新学習指導要領では「理科が『役に立つ』, 『楽しい』との回答が国際平均より低く, 理科の好きな子供が少ない状況を改善する必要がある」ことが具体的な改善事項として挙げられた (文部科学省, 2018). このことから, 理科の有用性を実感させる取り組みは, 我が国の理科

教育全体の課題として認識されているといえる.

大規模な学力調査で実施された質問紙における有用性の測定対象は, 以下のように2つに大別できる. 理科の有用性に該当する項目を抽出した濱保・山崎・岡田 (2019) によると, 国際数学・理科教育動向調査 (TIMSS2015; 国立教育政策研究所, 2017) では「理科を勉強すると日常生活に役に立つ」, 「将来, 自分が望む仕事につくために, 理科で良い成績をとる必要がある」の2項目が, 全国学力・学習状況調査 (国立教育政策研究所, 2018) では「理科の勉強は大切だと思いますか」, 「理科の授業で学習したことを普段の生活の中で利用できないか考えていますか」, 「理科の授業で学習したことは, 将来, 社会に出たときに役に立つと思いますか」, 「将来, 理科や科学技術に関する職業に就きたいと思いますか」の4項目がこれに該当するとしている. これらの項目表現より“活用”の対象となっている測定範囲は“日常生活”と“理系の職業選択”が想定されているといえる.

理科の有用性と理科の活用志向性との間の概念的差異は, これまで“有用性”として大きく括られていた概念を分解し, “有用性の肯定”と“主体性”という2つの段階を想定することで整理できる. 前述の質問紙で使用されている「日常生活に役に立つ」, 「社会に出たときに役に立つと思いますか」などの項目表現は, 生徒が理科の学習内容や考え方が日常生活において有用であることを認めるか否かを測定している. この水準は, 例えば日常の問題を解決する際に有用である可能性や, 理科の知識を保有していることが生活を豊かにしうることを認めるという水準であり, 自ら進んで理科の学習内容や考え方を生かすというような主体性の有無は問わない. それに対し, 「普段の生活の中で利用できないか考えていますか」, 「将来, 理科や科学技術に関する職業に就きたいと思いますか」などの項目では, 能動的に日常生活や職業選択に活用しようとする主体性が測定対象となっており, 理科の活用志向性はこうした心的状態を指す概念として位置づけられる. 有用性を肯定できない知識や考え方を活用するとは考えにくいことから, 有用性の肯定は活用志向性が存在するための必要条件であろう. 実際, 質問紙調査の結果を概観すると, 有用性の肯定の測定項目は, ほぼ一貫して活用志向性を問う質問項目よりも肯定的回答の割合が高い (e.g., 濱保ら, 2019; 国立教育政策研究所, 2017, 2018). この結果は, 活用志向性を問う

項目が有用性の肯定を問う項目よりも閾値（困難度）が高いことを示唆しており、この概念の枠組みを支持するものといえる。

こうした背景を受け、本研究では理科の活用志向性を、将来にわたり理科の学習内容や考え方を日常生活や理系の職業選択に活用しようとする主体的な態度と定義する。本研究では、日常生活における理科の活用志向性（以下、日常活用志向）と理系の職業選択に関する理科の活用志向性（以下、職業活用志向）を測定する2つの下位尺度を作成することを目指す。

3. 心理学の動機づけ理論と“理科の活用志向性”の比較を通じた概念規定

理科の活用志向性は、学習内容や考え方を主体的に日常生活や職業選択に活用しようとする態度である。すなわち、行為の生起に関する心理的メカニズムであることから、動機づけの一形態といえる。以下では、理科の活用志向性と類似した動機づけ概念を内包する理論である期待－価値理論（e.g., Eccles & Wigfield, 2002）、学習動機の2要因モデル（市川, 1995）との比較を通じ、類似性と概念的差異を整理する。

期待－価値理論において課題の選択や遂行は主観的な成功見込みである成功期待（expectancy of success）と、当該課題に取り組みたいと思わせる価値的側面である課題価値（task value）から説明される（e.g., Eccles & Wigfield, 2002）。課題価値は多面的な概念であり、活用志向性はそのうちの利用価値（utility value）に最も近いと考えられる。利用価値とは、狭義にはキャリア上の有用性を指す概念であったが、近年では日常生活の中での有用性を含めた解釈がなされる傾向がある（レビューとして、解良・中谷, 2019）。実際、前者は制度的利用価値、後者は実践の利用価値として細分化され（伊田, 2001）、それぞれを弁別して測定する下位尺度が作成されている（解良・中谷, 2014）。

実践的利用価値、制度的利用価値の概念範囲は、それぞれ理科の活用志向性のうち日常活用志向、職業活用志向と重複している。しかし、2つの利用価値が“有用性”を表す概念であるのに対し、前述したように理科の活用志向性は“有用性の肯定”に加えて生徒の“主体性”の有無を問題にする概念である。この点で両概念は異なるといえ、理科の活用志向性の方が利用価値よりも高度な水準であるといえる。

市川（1995）による学習動機の2要因モデルでは、

「賞罰の直接性」と「学習内容の重要性」の直交する2要因から動機づけを構造化している。両方の要因とも高い学習動機は「実用志向」と命名されており、「仕事や生活に生かす」ために勉強するという動機づけを表す。学習の目的が生活と仕事であるという点で、理科の活用志向性と大きく重複すると考えられる。

理科の活用志向性と実用志向は以下の点で相違がある。学習動機は「一般に、人はなぜ勉強しているのだと思いますか」、「あなた自身は、なぜ勉強していたのですか」という質問に対する自由記述よりモデル化された経緯がある（市川, 1995）。すなわち学習動機では学習に対する理由づけを問題にしており（鹿毛, 2013）、実用志向は「仕事や生活に生かす」ことが理科を勉強する理由である状態を表す。しかし、理科の活用志向性は将来にわたって理科を活用しようとする態度そのものに注目しており、勉強する理由を問題にするのではない。また、実用志向は「仕事や生活に生かす」ことを概念範囲に含んでいるが、理科の活用志向性を測定する場合、活用の対象を日常生活と職業選択に細分化して捉えることが望ましい。さらに、実用志向で想定されている「仕事」に生かすという理由づけは、職業選択に生かすという意味と、生産活動を最適化したり労働環境を整えたりするために生かすという職種によらない意味の2通りの解釈ができる。そのため、理科の活用志向性に特化した測定を行う際には、解釈の多義性を回避する工夫が必要だろう。

このように理科の活用志向性には類似する動機づけ概念が多くみられるものの、完全に同一の定義を採用している理論的枠組みはない。理科の活用志向性がSSH事業の趣旨や理科教育的意義に適合する概念であると考えられることから、新たにこれを測定できる尺度を開発する必要がある。

4. 本研究の目的

これまでに述べた背景を受け、本研究はSSH指定校の取り組みの評価を見据えた理科の活用志向性尺度を作成し、その信頼性と妥当性を検討することを目的とする。

II. 研究1：理科の活用志向性尺度の項目作成と因子構造、測定値の特徴の検討

1. 項目の作成

SSH指定校の取り組みの評価に活用できる理科の活

用志向性尺度を作成するためには、項目の反応閾値と項目表現の抽象度について工夫が必要である。

SSH 指定校には、理科の活用志向性が高い生徒が多く在籍していることが予想される。そうであるなら、多くの生徒が最高段階 (e.g., 5 件法における“5”) に反応するような閾値の低い項目のみで尺度構成をした場合、SSH 指定校では天井効果が生じ、理科の活用志向性の水準や変化幅を適切に評価できない。

同時に、SSH 指定校の第 1 学年 (以下、1 年生) では在校生の 80~100% を SSH の主対象としたカリキュラム構成をしているものの、第 2 学年 (以下、2 年生) や第 3 学年 (以下、3 年生) の SSH 主対象生徒は 10~20% の学校が多いことを考慮する必要がある (小林ら, 2015)。つまり、SSH 指定校の取り組みの評価には、全生徒を対象とした SSH 教育を行う 1 年生の教育評価と、全生徒の 20% 程度である SSH 主対象のクラスの 2, 3 年生の教育評価の 2 つが想定される。そのため、閾値の高い質問項目のみで尺度構成をすると、将来的に SSH 主対象でないクラス (文系など) に進級する生徒も多い 1 年生では床効果が生じうる。

また、SSH 指定校の教育効果を適切に評価するためには、一般校や SSH 主対象生徒でない生徒と比較を行うことが望まれていることから (小林ら, 2015)、理科の活用志向性尺度は一般校や SSH 主対象生徒以外の生徒にも使用されることを想定する必要がある。たとえ SSH 指定校の 1 年生で床効果が生じなくとも、一般校の 1 年生で床効果が生じるならば、通常の統計的分析では効果量を過小評価しうる。したがって、SSH 指定校の取り組みの評価を見据えた理科の活用志向性尺度は、基本的には① SSH 指定校および一般校の 1 年生を対象とした測定で床効果が生じず、② SSH 主対象生徒の 2, 3 年生を対象とした測定で天井効果が生じないことの 2 つの条件を満たす必要がある。

しかし職業活用志向では、これが高い生徒は 2, 3 年生で理系クラスに進級すると考えられることから、2, 3 年生の文系クラスにおいて職業活用志向の下位尺度で床効果が発生することは自然であると考えられる。そのため、職業活用志向における文系クラスでの床効果は構成概念の定義上、許容されるだろう。一方、日常活用志向は進路選択と直結しない変数であるため、2, 3 年生の文系クラスは理系クラスよりも測定値が低いと予想されるものの、それでも床効果が生じないよう項目表現に配慮する必要がある。

次に、項目表現の抽象度について論じる。教育現場での尺度の活用を見据える場合、項目数が多すぎると授業時間を圧迫する恐れがあることから、質問項目は可能な限り少なくすることが望まれる。しかし、測定したい構成概念が広義であるにも関わらず少数の具体的項目で構成概念の諸側面を捉えようとするとは“構成概念の代表性不足”に陥る恐れがある (平井, 2010)。極めて少ない質問項目で構成概念を測定する必要がある場合、構成概念それ自体を反映する非常に抽象度の高い項目表現により当該構成概念の諸側面をまとめて表現する“包括的単一項目 (comprehensive single items)”という方法を用いることが考えられる (Konstabel, Lönnqvist, Walkowitz, Konstabel, & Verkasalo, 2012)。これと類似した考え方に基づいて作成されたごく少数の項目を用いた尺度においても、一定の妥当性が認められている (e.g., Konstabel et al., 2012; 箕浦・成田, 2013; Núñez-Peña, Guilera, & Suárez-Pellicioni, 2014; Robins, Hendin, & Trzesniewski, 2001)。一方で、項目表現を抽象的にしすぎると、項目の解釈に個人差が生じ、生徒によって異なる意味を見出ししてしまう可能性が生じる (村上, 2006)。例えば「理科を日常生活で活用したいですか」という抽象度の高い 1 項目のみで測定すると、想起する日常生活場面が個人によって異なる可能性が高く、測定対象が曖昧になる。

こうしたことを考慮し、本研究では、日常活用志向 6 項目、職業活用志向 6 項目で測定できるよう設計し、項目の抽象度について以下のような配慮をする。日常活用志向では、具体的な日常生活場面 4 項目 (ニュース, 教育, 勤労, 消費行動) と抽象度の高い 2 項目を混在させることで、解釈の多様性を回避しながら少数項目で構成概念全体を捉えることを目指す。

一方で、職業活用志向は、全項目で具体的な職業名を挙げ、それに就きたい程度を尋ねると、以下のような測定の妥当性が侵害される問題が生じる。例えば、希望職種がある 1 つの理系の職業に定まっている生徒の回答は、該当する 1 項目のみ最高段階に反応するだろう。また、希望する職業が項目に含まれていない場合、たとえ理系の職業を志望していても測定値は最小値になりえる。こうしたことは構成概念の定義上、測定の妥当性が担保されていると認められない。そこで、理系の職業全般に当てはまる特徴を抽象的に表現する 5 項目と、具体的な職業名 (科学者) の 1 項目から尺度を構成し、想起する具体的職業が個人によって異なる

るとしても、理系の職業を志向する程度が尺度得点に反映されるようにする。

以上のことに配慮し、理科教育学研究者2名（1名は元高校理科教員）の協議によって、日常活用志向6項目、職業活用志向6項目を作成した（表2参照）。

2. 方法

a. 調査対象者

協力校は、学力水準を統制するため、高校入試偏差値と大学進学者数が同程度の、SSH指定校1校（A校）の1, 2, 3年生とSSHに指定されていない一般校1校（B校）の1, 2年生を対象とした。

A校は、1年生に対しては全生徒を対象にSSHのカリキュラムを実施していた。また、2, 3年生に対してはSSH指定クラスの生徒（以下、SSH主対象生徒）に対し、SSH独自のカリキュラムに基づく教育を行っていた。

b. 測定変数

理科の活用志向性尺度 作成した日常活用志向（6項目）、職業活用志向（6項目）を使用した。回答は「1. 全く当てはまらない」、「2. 当てはまらない」、「3. どちらともいえない」、「4. 当てはまる」、「5. よく当てはまる」までの5件法で求めた。

c. 手続き

調査は2019年1月に各協力校の理科教員によって授業中に実施された。協力校の事情を勘案し、B校の3年生は対象から除外した。生徒には、調査結果が成績に影響しないこと、個人情報には必ず守られ、また出席番号など個人を識別できる情報を記載する必要はないことを説明した。

3. 分析の方針

本研究では心理尺度の開発を目的としているため、信頼性と妥当性の検討が必要である。

本尺度の信頼性は内的整合性に加えて、テスト情報量曲線（Test Information Curve；以下、TIC）によって検証する。内的整合性による信頼性の推定は、1つの検討対象の尺度について1つの指標しか得られないが、TICは尺度で測定される個人特性値（ θ ）に応じた情報量（測定精度）を検討できる。本尺度はSSH指定校の取り組みの評価を見据えているため、平均値より高い層に対しても測定精度が担保されている必要がある。

研究1では尺度の妥当性を所属クラス（グループ）の平均値と平均値差によって検討する。平均値については、文系クラスにおける職業活用志向以外に天井効果（ $Mean + SD > 5.00$ ）および床効果（ $Mean - SD < 1.00$ ）が生じないことを確認する。

またグループごとの比較では、以下のような平均値差がみられると予想される。SSH主対象生徒は、SSHのカリキュラムにより理科学習への動機づけを高められていることに加え、そもそもSSH指定クラスを志望する生徒は初期の時点より動機づけが高いと考えられることから、SSHの主対象でない通常の理系クラスや文系クラスの生徒より、日常活用志向、職業活用志向とも高い値だろう。また、高校では生徒の希望進路に応じて文系や理系クラスに分けられることから、2, 3年生の所属クラスの決定過程には個人の職業活用志向が直接的に関与すると考えられることから、職業活用志向のグループ間差は、日常活用志向よりも大きな効果量であるだろう。

さらに、SSH指定校の取り組みを評価する場合において、一般校の生徒などとの比較が推奨されていることから（小林ら、2015）、所属する学校およびクラスにおける特異項目機能（differential item functioning：以下、DIF）の検討が必要である。DIFとは個人特性値が等しいにも関わらず所属グループによって回答傾向に差異がある状態である。DIFが存在する場合、当該項目にグループ間の平均値差を見出したとしても、測定対象となる構成概念の差が反映されたものであるのか、それともバイアスが反映されたものであるのか不透明となる（田崎、2008）。そのため、研究1ではDIFの検討を行い、SSH指定校と一般校、およびSSH主対象生徒とそれ以外の生徒での尺度の等価性を確認し、本尺度が群間比較に耐えうるものであることを確認する。

4. 結果

a. 分析対象者

研究の同意と回答が得られた1021名（男子449名、女子557名、無回答15名）を分析対象とした（表1）。両学校とも、1年生は文系・理系クラスに分かれていなかった。またA校の2年生33名、3年生41名がSSH主対象生徒であるSSH指定クラスに所属していた。

以降の因子分析はロバスト最尤法によって行い、欠損値推定の方法には完全情報最尤法を用いた。

b. 理科の活用志向性尺度に対する因子分析

理科の活用志向性尺度に対して探索的因子分析を行った。固有値の減衰状況は、6.96, 1.76, 0.53...であった。カイザー基準、平行分析（サンプリング数100；図1）、最小平均偏相関（Minimum Average Partial；以下、MAP；0.95, 0.27, 0.38...）で2因子構造が提案された。また、2因子における累積寄与率は72.62%であり、十分な大きさと判断された。以上のことから、理科の活用志向性尺度は想定通り2因子構造であると判断した。

各下位尺度と項目の特徴を分析するため、Mplus（Ver 8；Muthén & Muthén, 2017）を用いてカテゴリカル因子分析（ie. 項目反応理論における段階反応モデル）を行い、因子負荷量と閾値を求めた。分析モデルは確証的因子分析モデルであり、日常活用志向の6項目に共通する潜在変数と、職業活用志向の6項目に共通する潜在変数の2因子を設定し、交差負荷（cross-loading）は認めなかった。また、因子間には相関を仮定した。

分析の結果を表2に示した。モデルの適合度は許容範囲内であった（CFI = 0.986, TLI = 0.983, RMSEA =

0.093)¹⁾。因子負荷量は抽象的な表現を採用した項目も具体的な表現を採用した項目もほぼ同程度に高かった。

c. DIFの検討

DIFの検討は、SSH指定校と一般校の1年生を比較することと、2, 3年生においてSSH主対象生徒とそれ以外の生徒で比較することを見据え、それぞれに対して分析した。検証方法は、カテゴリカル因子分析モデルの多母集団同時分析による情報量規準の比較と、熊谷（2012）による“指標K”によって行った。

1年生では、所属学校（A校 or B校）をグループ変数とした。また2, 3年生ではSSH主対象生徒とそうでない生徒（ie., 文系・理系）をグループ変数とした。

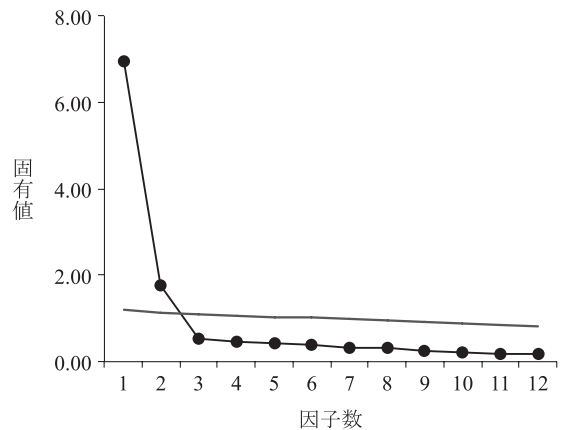


図1 理科の活用志向性尺度の因子数と固有値、および平行分析の結果

表1 本研究の分析対象者の内訳

	文理なし		文系		理系		SSH指定		合計
	A校	B校	A校	B校	A校	B校	A校	B校	
1年生	200	264	0	0	0	0	0	0	464
2年生	0	0	38	185	41	72	33	0	369
3年生	0	0	85	0	62	0	41	0	188
合計	200	264	123	185	103	72	74	0	1021

表2 理科の活用志向性尺度の因子分析結果（カテゴリカル因子分析）

項目	具体 / 抽象	因子負荷量		閾値			
		推定値	(SE)	1	2	3	4
<日常生活志向> α = .908							
1 学校を卒業した後も、理科で学んだ内容を日常生活に生かしたい	抽象	.847	(0.015)	-1.104	-0.529	0.077	1.087
2 科学に関するニュースを見たとき、理科で学んだ知識と関連づけて考えたい	具体	.834	(0.016)	-1.164	-0.465	0.189	1.197
3 もし将来、子どもの教育に関わることがあれば、理科の知識や考え方を教えたい	具体	.779	(0.019)	-1.007	-0.451	0.296	1.160
4 理系の職業ではなくとも、理科で学んだことを仕事の場面で生かしたい	具体	.884	(0.013)	-1.099	-0.537	0.124	1.004
5 理科で学んだ学習内容を、積極的に生活に役立てたい	抽象	.908	(0.011)	-1.369	-0.720	0.062	1.026
6 商品の広告を見て買うか買わないかを決めるとき、理科の知識を判断材料にしたい	具体	.787	(0.019)	-1.014	-0.489	0.341	1.216
<職業活用志向> α = .935							
7 将来は、理科で学んだ内容を使う仕事に就きたい	抽象	.861	(0.014)	-0.589	0.136	0.697	1.252
8 実験や観察によって、まだわかっていないことを調べる仕事がいい	抽象	.929	(0.008)	-0.546	0.254	0.904	1.549
9 理系の大学等に進んで、将来は、そこで研究したことを生かした仕事に就きたい	抽象	.928	(0.009)	-0.354	0.276	0.820	1.350
10 自分は、理系分野の“科学者”になりたい	具体	.904	(0.011)	0.067	0.679	1.477	1.917
11 理科の知識を使って、何か新しいことを発見するような職業に就きたい	抽象	.935	(0.009)	-0.324	0.356	0.987	1.588
12 理科で学んだ考え方を生かして、何かを開発するような仕事でみたい	抽象	.886	(0.013)	-0.409	0.236	0.792	1.478
因子間相関		.690 (0.023)					

まず、田崎 (2007)、船越・田崎・潮村 (2013) による手順を参考に、表 2 と同様のカテゴリカル因子分析モデルについて多母集団同時分析を行った。各グループで全項目の因子負荷量と閾値に等値制約を設定したモデルをベースモデルとし、情報量規準 (AIC, BIC) を求めた。次にベースモデルを基礎に、1 項目ずつ当該項目のみ因子負荷量と閾値を自由推定させたときの情報量規準を求めた。また、各下位尺度について “Easy DIF” (熊谷, 2012) を用いて指標 K を求めた。

結果を表 3 に示した。いくつかの項目では AIC がベースモデルよりわずかに自由推定モデルを支持したが、BIC は一貫して等値制約を設定したベースモデルを支持した。また指標 K は一貫して DIF を検出しなかった。以上のことから、SSH 指定校と一般校の 1 年生の間の DIF は検出されなかったと判断した。

同様に、2, 3 年生を対象とした分析の結果を表 3 に示した。いくつかの項目では AIC がベースモデルよりわずかに自由推定モデルを支持したが、BIC は一貫して等値制約を設定したベースモデルを支持した。また指標 K は一貫して DIF を検出しなかった。以上のことから、SSH 主対象の生徒とそうでない生徒の間の DIF は検出されなかったと判断した。

d. 理科の活用志向性の信頼性の検討

各下位尺度の信頼性を検討した。日常活用志向の内的整合性は十分に高く ($\alpha = .908$)、TIC を参照すると、個人特性値 θ に対して、およそ $-1.5 < \theta < 1.5$ の範囲で高い値を示していた (図 2)。職業活用志向の内的整合性も十分に高く ($\alpha = .935$)、TIC を参照すると、およそ $-0.5 < \theta < 1.7$ の範囲で高い値を示していた (図 2)。

e. グループごとの測定値の比較

以降では、各グループの測定値の特徴を検討する。分析には学校現場での活用を見据え、項目平均値による尺度得点を用いた (得点幅: 1.00~5.00)。カテゴリカル因子分析で推定した個人特性値 θ と尺度得点の相関係数を求めたところ、極めて強い相関が確認されたことから、(日常活用志向: $r = .985, p < .001$, 職業活用志向: $r = .967, p < .001$) 尺度得点による指標でも大きな問題はないと判断した。

記述統計量を表 4 に示した。記述統計量を参照すると、2, 3 年生の文系クラスの職業活用志向で床効果がみられたが、その他のグループでは天井効果や床効果はみられなかった。

この結果は当初の目的通りであり、SSH 指定校の教育効果を適切に評価できること、および一般校や SSH

表 3 1 年生および 2, 3 年生における DIF の分析結果

項目	1 年生 (SSH 指定校 vs 一般校)			2,3 年生 (文系・理系 vs SSH 指定クラス)		
	多母集団同時分析		指標 K	多母集団同時分析		指標 K
	AIC	BIC		AIC	BIC	
ベースモデル	12703.776	12975.425		14863.482	15148.771	
日常活用志向						
1	12704.256	12996.485	0.171	14863.821	15170.723	0.134
2	12703.558	12995.786	0.178	14869.273	15176.175	0.134
3	12703.068	12995.297	0.139	14864.345	15171.247	0.077
4	12709.804	13002.033	0.128	14870.769	15177.672	0.147
5	12705.076	12997.305	0.085	14871.801	15178.703	0.060
6	12707.420	12999.648	0.151	14870.951	15177.853	0.096
職業活用志向						
7	12699.477	12991.705	0.244	14862.509	15169.412	0.351
8	12707.761	12999.989	0.131	14869.030	15175.932	0.138
9	12706.008	12998.237	0.151	14856.470	15163.372	0.083
10	12703.117	12995.345	0.141	14867.567	15174.469	0.058
11	12708.418	13000.646	0.089	14867.125	15174.027	0.067
12	12706.666	12998.894	0.123	14863.133	15170.036	0.148

注) ベースモデルよりも小さい値を示した情報量規準について、太字で示した。指標 K は (カテゴリ数 - 1) × 0.1 が DIF 検出の基準とされている (熊谷, 2012)。本研究では、指標 K > 0.40 をもって DIF が検出されたかと判断する。

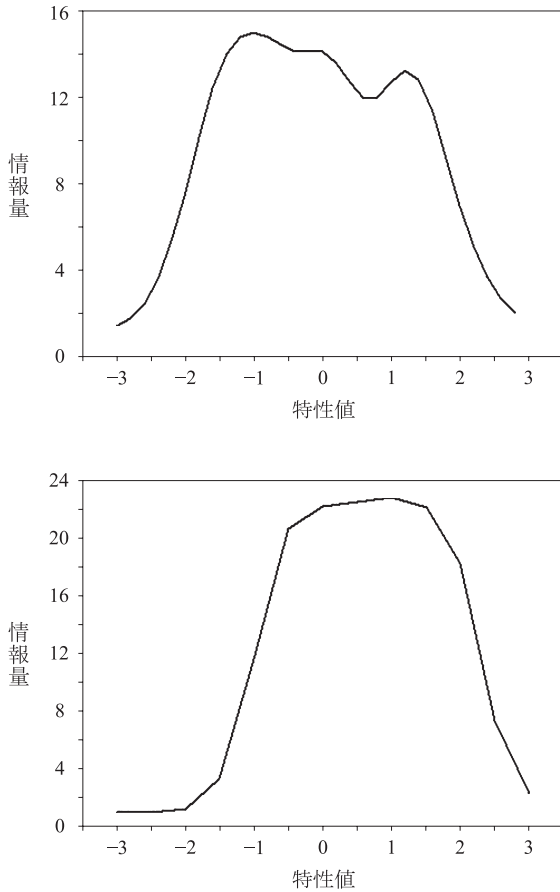


図2 日常活用志向（上）と職業活用志向（下）のTIC

主対象生徒でない生徒との群間比較が可能であることを示唆する。

1年生におけるSSH指定校と一般校の平均値差を検討するため、Welchのt検定を行った。日常活用志向はA校の方がB校よりも高く ($t(441.177) = 5.853, p < .001, d = 0.541 [0.351, 0.731]$)、職業活用志向も同様であった ($t(398.561) = 3.475, p < .001, d = 0.333 [0.145, 0.520]$)。

次に、2, 3年生における文系・理系・SSH指定クラスの差を検討するため、分散分析を行った。多重比較はHolm法によった。日常活用志向の群間差は有意であり ($F(2,551) = 60.090, p < .001, \eta^2 = .179 [0.124, .233]$)、文系クラスより理系クラス ($d = 0.727 [0.536, 0.919], p < .001$)、SSH指定クラス ($d = 1.245 [1.043, 1.446], p < .001$) が高く、理系クラスよりSSH指定クラスの方が高かった ($d = 0.517 [0.329, 0.706], p < .001$)。

表4 各グループの記述統計量

グループ	日常活用志向			職業活用志向		
	n	Mean	(SD)	n	Mean	(SD)
対象者全体	1006	3.158	(0.981)	1006	2.295	(1.026)
1年生						
A校 (SSH指定校)	190	3.463	(0.765)	190	2.455	(0.951)
B校 (一般校)	262	2.999	(0.917)	262	2.145	(0.917)
2, 3年生						
文系クラス	307	2.752	(1.063)	307	1.762	(0.835)
理系クラス	174	3.448	(0.784)	173	2.815	(0.873)
SSH指定クラス	73	3.943	(0.839)	74	3.408	(1.086)

職業活用志向の群間差は有意であり ($F(2,551) = 144.473, p < .001, \eta^2 = .344 [0.282, .399]$)、文系クラスより理系クラス ($d = 1.190 [0.990, 1.391], p < .001$)、SSH指定クラス ($d = 1.860 [1.640, 2.080], p < .001$) が高く、理系クラスよりSSH指定クラスの方が高かった ($d = 0.670 [0.479, 0.860], p < .001$)。

5. 研究1の考察

理科の活用志向性尺度に対する因子分析の結果、当初の目的通り日常活用志向と職業活用志向の2つの因子が確認され、適合度は許容できる水準であった。この結果は本尺度の因子的妥当性を支持するものである。

DIFの検討を行ったところ、SSH指定校と一般校の1年生の間、および2, 3年生のSSH主対象生徒とそうでない生徒の間にDIFは存在しないことが示唆された。この結果は、当該群間において項目機能は等価であり、群間比較に耐えうることを示唆する。

本尺度の信頼性を検討したところ、 α 係数は十分な値であり、またTICは θ が1.00より大きい範囲まで高い値を示した。この結果は、日常活用志向では θ が1.5程度まで、職業活用志向では1.7程度まで高い精度で測定できることを示しており、本尺度が幅広い特性値の生徒に対して高い信頼性を有していることを示唆するものである。

本尺度の測定値を検討したところ、当初の目的通り、文系クラスの職業活用志向以外に天井効果、床効果は確認されなかった。この結果は、SSH指定校の生徒のような高い理科の活用志向性を持っていると考えられる生徒であっても、適切に測定することが可能であることを示唆する。またグループ間比較では、当初の仮説通り、1年生では一般校 < SSH指定校であった。また、2, 3年生では文系クラス < 理系クラス < SSH指定クラスであり、差の大きさの程度を示す効果量は

職業活用志向の方が日常活用志向よりも大きかった。この結果は、本尺度の測定値は現実の実態をよく捉えられていることを示唆するものである。

Ⅲ. 研究2：理科の活用志向性尺度と関連が予想される心理変数との相関分析による妥当性の検討

1. 序論

研究1では作成した理科の活用志向性尺度について、 α 係数とTICによる信頼性の確認、因子分析と群間比較による妥当性の確認が行われた。しかし、概念間の類似性や理論的背景より事前に相関があると予想される心理変数との相関分析を通じた妥当性の検討は行われていない。そのため研究2では相関分析を通じた妥当性の検討を行う。

使用する変数は、第1章3節で述べた期待一価理論における利用価値の概念である。前述したように実践的利用価値、制度的利用価値の概念範囲はそれぞれ日常活用志向、職業活用志向と重複している。そのため、日常活用志向は実践的利用価値と正の相関があり、かつその相関は制度的利用価値との相関よりも強いと予想される。対して、職業活用志向は制度的利用価値と正の相関があり、かつその相関は実践的利用価値との相関よりも強いと予想される。研究2では、この仮説を相関の差の検定を通して検討する。また興味価値と理科の活用志向性との相関に関しては、正の相関があるとの予測はできるものの(eg, 解良・中谷, 2014), 具体的な相関の程度については不透明である。そのため、これを探索的に検討する。

関連が予想される心理変数との相関の有無は当該尺度の構成概念妥当性を確認する上で重要であるが、極端に強い相関を示すのであれば、用いた心理尺度と実質的に同一の構成概念を捉えていることを示唆する。その場合、新たな測定尺度を開発する意義は小さい。これを考慮し、本研究では全分散のうち約半分を説明する $r = .700$ ($r^2 = .490$)を超えた場合、高度に類似した構成概念を捉えていると判断する。

ところで、研究2では研究1と異なる対象者に調査を行うため、理科の活用志向性尺度について再度因子構造の検討が可能である。もし研究1と同様の因子構造が再現されたならば、本尺度の因子構造には再現性があり、因子的妥当性は頑健であるといえる。そこで、研究2では研究1のサンプルと合わせた多母集団同時分析を通して、因子構造の再現性を確認する。

2. 方法

a. 調査対象者

SSH指定校であるC校の1, 2, 3年生を調査対象とした($n = 795$)。

b. 測定変数

尺度への回答は研究1と同様の5件法で求めた。なお使用した項目はすべて附録に示した。

理科の活用志向性尺度 日常活用志向(6項目)、職業活用志向(6項目)を使用した。

理科の課題価値評価尺度 解良・中谷(2014)が作成した尺度より、興味価値(4項目: “理科の勉強の内容は、面白いと思います”, “理科の勉強は、楽しいと思います”, “理科で勉強する内容は、つまらないと思います”(逆転項目), “理科で勉強する内容に、興味があると思います”), 実践的利用価値(3項目: “理科の内容は、私の身の回りで役に立っていると思います”, “理科を勉強することで、身の回りのできごとや現象のしくみを理解する事ができると思います”, “理科の内容をよく知っていると、ふだんの生活の中で役に立つことがあると思います”), 制度的利用価値(3項目: “理科を勉強することは、希望の進路を実現するために特に大事だと思います”, “理科の勉強ができることは、就職するときに役に立つと思います”, “理科の勉強は、大学入試で合格するために特に重要だと思います”)を使用した。

c. 手続き

調査は2020年7月に協力校の理科教員によって授業中に実施され、研究1と同様の手続きによった。

3. 結果

a. 分析対象者

795名(男子413名, 女子381名, 無回答1名)全員を分析対象とした。

b. 理科の活用志向性尺度に対する因子分析と因子構造の再現性の検討

研究2のデータのみを用い、理科の活用志向性尺度に対して探索的因子分析を行った。固有値の減衰状況は、6.85, 1.81, 0.65...であった。カイザー基準、平行分析、MAP(0.10, 0.03, 0.04...)によって2因子構造が提案された。2因子における累積寄与率は72.19%であり、十分な大きさと判断された。以上のことから、研究1と同様に2因子構造であると判断した。

研究2のデータに対し、研究1と同様の確証的因子

分析（カテゴリカル因子分析）を行った。因子負荷量、閾値、因子間相関は研究1とはほぼ同程度であった（表5）。モデルの適合度は許容範囲内であった（CFI = 0.929, TLI = 0.911, RMSEA = 0.090）¹⁾。

続いて、研究1と研究2のデータを用いて、多母集団同時分析を行った。比較したモデルは因子構造のみが研究1と研究2の集団で等しい配置不変モデルと、配置不変モデルに加えて全ての因子負荷量、閾値、因子間相関が等しい測定不変モデルの2つであった。測定不変モデルの情報量規準（AIC = 50216.631, BIC = 50579.520）は、配置不変モデルの情報量規準（AIC = 50232.135, BIC = 50908.428）よりも小さかった。また測定不変モデルの適合度は良好であった（CFI = 0.936, TLI = 0.933, RMSEA = 0.079）¹⁾。このことから、因子構造と各パラメータの再現性が確認されたと判断した。

c. 課題価値との相関分析

理科の活用志向性尺度の妥当性を検討するため、相関が予想される変数として採用した各課題価値概念との相関分析を行った。

分析は、理科の活用志向性の2因子、各課題価値について潜在変数を抽出し、5つの潜在変数間の相関を想定した構造方程式モデリングによって行った²⁾。各課題価値の因子負荷量は、興味価値（.652 < β < .932, all ps < .001）、実践的利用価値（.717 < β < .826, all ps < .001）、制度的利用価値（.728 < β < .867, all ps < .001）の全てで十分な水準であった。

因子間相関の分析結果を図3に示した。本モデル

の適合度は良好であった（CFI = 0.930, TLI = 0.919, RMSEA = 0.065）¹⁾。相関係数の差を検討したところ、日常活用志向と実践的利用価値との間の正の相関は、日常活用志向と制度的利用価値との相関よりも強かった（z = 2.191, p = .028）。同様に、職業活用志向と制度的利用価値との間の正の相関は、職業活用志向と実践的利用価値との相関よりも強かった（z = 5.143, p < .001）。また、日常活用志向、職業活用志向とも、興味価値との正の相関が認められた。理科の活用志向性と各課題価値概念の相関係数の推定値は .700 を下回っていた。

4. 考察

理科の活用志向性尺度に対する因子構造の分析の結

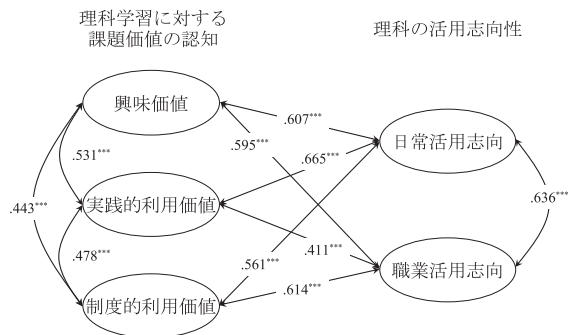


図3 構造方程式モデリングによる因子間の相関

注) 図示された矢印はすべて因子間相関を意味する。そのためすべて双方向の矢印で示している。また各項目は省略した。*** p < .001

表5 研究2のデータを用いた因子分析の結果（カテゴリカル因子分析）

項目	因子負荷量		閾値			
	推定値	(SE)	1	2	3	4
〈日常活用志向〉 α = .903						
1 学校を卒業した後でも、理科で学んだ内容を日常生活に生かしたい	.880	(0.015)	-1.274	-0.664	-0.065	0.999
2 科学に関するニュースを見たとき、理科で学んだ知識と関連づけて考えたい	.854	(0.019)	-1.277	-0.660	0.039	1.057
3 もし将来、子どもの教育に関わることがあれば、理科の知識や考え方を教えたい	.799	(0.020)	-1.167	-0.535	0.186	1.049
4 理系の職業ではなくとも、理科で学んだことを仕事の場面で生かしたい	.845	(0.017)	-1.268	-0.669	0.005	0.890
5 理科で学んだ学習内容を、積極的に生活に役立てたい	.900	(0.014)	-1.527	-0.757	-0.051	0.930
6 商品の広告を見て買うか買わないかを決めるとき、理科の知識を判断材料にしたい	.714	(0.027)	-1.221	-0.606	0.292	1.175
〈職業活用志向〉 α = .936						
1 将来は、理科で学んだ内容を使う仕事に就きたい	.815	(0.020)	-0.769	-0.078	0.521	0.998
2 実験や観察によって、まだわかっていないことを調べる仕事にしたい	.907	(0.012)	-0.600	0.131	0.852	1.508
3 理系の大学等に進んで、将来は、そこで研究したことを生かした仕事に就きたい	.915	(0.011)	-0.527	0.139	0.689	1.280
4 自分は、理系分野の“科学者”になりたい	.923	(0.010)	-0.115	0.549	1.284	1.874
5 理科の知識を使って、何か新しいことを発見するような職業に就きたい	.959	(0.008)	-0.386	0.296	0.921	1.484
6 理科で学んだ考え方を生かして、何かを開発するような仕事にしてみたい	.904	(0.013)	-0.494	0.200	0.774	1.370
因子間相関	.639	(0.030)				

果、研究1と同一の因子構造、因子負荷量、閾値、因子間相関を想定した測定不変モデルが指示された。この結果は、本研究で示した因子構造の再現性を確認したものであり、十分な因子的妥当性を有していることを示すものと考えられる。

理科の活用志向性と各課題価値概念との相関分析の結果、当初の仮説通りの相関のパターンが確認された。また相関係数の値は極端に大きいものではなかった。この結果は、本研究で作成された理科の活用志向性尺度が構成概念妥当性の中でも外的な側面の証拠(村山, 2012)を有することと、また既存の測定尺度とは異なる構成概念を測定対象とすることを示すものであると考えられる。

IV. 総合考察

本研究の目的は、SSH指定校の取り組みの評価を見据えた理科の活用志向性尺度を作成し、その信頼性と妥当性を検討することであった。本尺度では日常活用志向と職業活用志向の2つの構成概念を測定する下位尺度を作成した。以降では得られた結果をもとに、本研究の目的が達成されたといえるかについて考察する。

1. 理科の活用志向性尺度の因子構造

研究1、研究2でそれぞれ因子分析を行った結果、当初の想定通りの2因子構造が採択され、適合度も許容範囲内であった。これらの結果は、本研究で測定対象とした日常活用志向と職業活用志向の2つの下位概念を測定できたことを示すと考えられる。

また、学校現場での活用を見据えて、項目の抽象度を操作し、少数の項目で構成概念を捉えることを目指した。分析の結果、抽象的な項目も具体的な項目も同程度に因子負荷量が高かったことから、それぞれが同じ構成概念を捉えられていると考えられる。

さらに、研究2では因子分析結果の再現性を確認するため、多母集団同時分析によってモデル比較を行ったところ、厳しい等値制約を設けた測定不変モデルが採択された。以上の結果を総括すると、本研究で測定対象とした日常活用志向と職業活用志向の2つの下位概念を測定できたものと考えられ、また十分な因子的妥当性が確認されたといえる。

2. 理科の活用志向性尺度の信頼性

作成した理科の活用志向性尺度の内的整合性は研究

1、研究2ともに十分な水準であった。また、SSH指定校には理科の活用志向性が高い生徒が多く在籍していると考えられるため、特に高い特性値を持つ生徒の測定精度が高く保たれている必要があった。TICを求めたところ、分布はやや右側に寄っており、日常活用志向では θ が1.5程度まで、職業活用志向では1.7程度まで高い精度で測定できることが明らかになった。正規分布の累積分布関数を参照すると、 $\theta > 1.5$ の生徒の割合は6.68%程度と考えられることから、理科の活用志向性尺度の得点分布が正規分布すると仮定すると、本尺度はほとんどの生徒の理科の活用志向性について高い精度で測定できるといえる。さらに両尺度とも、TICを参照すると θ が負の領域まで高い測定精度を持つことも明らかになった。そのため、SSHの主対象者でない生徒や一般校の生徒に対しても一定の測定精度が保証できるといえ、本尺度の適用範囲は広いといえる。

3. 理科の活用志向性尺度の妥当性

本尺度の妥当性について、研究1ではDIFの検討、尺度得点の天井効果と床効果の確認、グループ間の平均値差の側面から検討した。

SSH指定校の取り組みの評価には測定値の比較が推奨されていることを受け(小林ら, 2015)、本研究では1年生についてはSSH指定校と一般校、2、3年生についてはSSH主対象生徒とそれ以外の生徒でDIFを検討した。分析の結果、DIFは検出されなかった。この結果は、これらの群間では項目機能の等価性が確認されたことを意味しており、群間比較に耐えうる尺度であるといえる。

測定値を確認したところ、文系クラスにおける職業活用志向に床効果がみられた以外の天井効果と床効果は確認されなかった。文系・理系のクラス選択は生徒の希望進路に応じて決定されることから、文系クラスの職業活用志向が極めて低い値を示すことは自然な現象であり、むしろこの結果は、本尺度の測定値が現実の実態をよく反映していることを示している。そのため、今回作成した理科の活用志向性尺度では、問題のある天井効果および床効果は生じていないといえ、測定上の妥当性は確認されたといえる。

この結果の実用上の意義は以下2点にある。1点目は、尺度得点による指標化でもあらゆるグループ(e.g., SSH指定校, SSH指定クラス)においても個人差の識別が可能なことである。すなわち、どの群においても、

全員が全項目に“5”と回答するなどの測定値の集中は生じず、分散が存在することから、理科の活用志向性の個人差や集団差を適切に評価できる。また、推定した個人特性値 θ と尺度得点との間に極めて強い正の相関が確認されたことから、教育現場にとって必ずしも容易でない統計的手法を用いずとも、教育的介入が必要なクラスや生徒個人を同定できる。そのため、本尺度は教師の授業改善に貢献できると考えられる。2点目はSSH指定校の取り組みの評価のために一般校などと尺度得点の比較をする際、打ち切り正規分布などの特殊な分布を想定しない単純な効果量の推定でも、過小評価せずに済むと考えられることである。そのため本尺度は、SSH事業の目的の1つである「旧来の学習指導要領の改善に資する試みの実施」（小林ら，2015）についての実証的資料を提供することに貢献できると考えられる。

理科の活用志向性尺度のグループ間の平均値差の分析では、事前の作業仮説の通り、日常活用志向、職業活用志向とも文系クラス<理系クラス<SSH指定クラスの順に値が大きく、また日常活用志向 ($\eta^2 = .179 [0.124, 0.233]$) よりも職業活用志向 ($\eta^2 = .344 [0.282, 0.399]$) の方が群間差の効果量が大きかった。この結果は構成概念の概念規定から予想される実態と整合するものである。したがって、本研究で作成した理科の活用志向性尺度の測定値は現実の実態をよく反映しているといえる。

SSH指定校と一般校の1年生を比較したところ、SSH指定校の1年生の方が日常活用志向、職業活用志向とも、およそ中程度の効果量で大きな値を示した。この結果は、SSH指定校のカリキュラムが理科の活用志向性への教育的効果を有していると考えられることや、そもそもSSH指定校を志望する生徒は一般校を志望する生徒よりも理科の動機づけが高いと考えられることと矛盾しない。本研究の結果からは、どちらの背景がより強く働いたかを明らかにすることはできないが、いずれにせよ本尺度の構成概念妥当性についての証拠を補強するものといえる。

研究2では、本尺度の妥当性について理論的に関連が予想される他の変数との相関分析を通して妥当性を検討した。本研究ではEccles & Wigfield (2002) の期待-価値理論における課題価値のうち2つの利用価値概念との類似性を検討した。事前の仮説通り、日常活用志向は実践的利用価値と強い相関を示し、職業活用

志向は制度的利用価値と強い相関を示した。この結果は本尺度が構成概念妥当性の中でも外的な側面の証拠(村山, 2012)を有していることを示すものである。また、各利用価値概念との相関係数の大きさは、実質的に同一の構成概念を捉えていることを考えられる水準ほど大きくなかった。この結果は、既存の測定尺度とは異なる概念範囲を測定対象としていることを示すものであり、本研究が提案する理科の活用志向性という構成概念の新規性を支持するものと考えられる。

4. 今後の課題

本研究は尺度の開発が目的であったため、実際にSSH指定校のカリキュラム・マネジメントに貢献した実績はまだない。そのため、実際にSSH指定校の教育評価に本尺度を活用するとともに、本尺度を導入したことで得られた利点を検証していくことが必要である。

これに関連して、SSH指定校の取り組みの評価に使用する場合の運用方法を検討してみたい。本研究のような1時点の調査では、SSH指定校と一般校で差を見出したとしても、SSHのカリキュラムによる教育効果であるのか、それとも単に入学者の初期値が高かったに過ぎないかを検討することができない。理想論を言えば、適切にこれを評価するために年間数回の調査を行い、潜在曲線モデル(Latent Curve Model)などを用いて、初期値(切片)と変化(傾き)を弁別して捉えることが望ましい。また、潜在曲線モデルの多母集団同時分析によって、SSH指定校の方が一般校よりも傾きの推定値が大きいために示されれば(非等値モデルの適合度が良いならば)、SSH指定校のカリキュラムによる理科の活用志向性への教育効果である可能性に言及でき、「旧来の学習指導要領の改善に資する試みの実施」(小林ら, 2015)についての実証的資料となるだろう。しかし、理想的にはこうした運用方法が良いだろうが、このような分析は学校現場で容易くできるものではないかもしれない。そのため、研究者は現場で可能な方法の中で、どのような運用方法を採用するのが良いのか、例えば、分析用プログラムの開発や尺度の標準化など、分析の妥当性を担保しつつ、実施可能な具体的な運用方法を示すことが必要だろう。

また、本研究では尺度の信頼性と妥当性について、全ての側面を実証できたわけではない。例えば、再検査信頼性、学力水準によるDIFなど、引き続き批判的な検討を続けることが重要であると考えられる。

注

- 1) Mplusにおける最尤法によるカテゴリカル因子分析では適合度指標が算出されない。そのため、同モデルをカテゴリカル因子分析ではなく通常の因子分析を行ったときの適合度指標を記載した。
- 2) 尺度得点（各項目に対する回答の平均値）を用いた相関分析では、尺度の信頼性が理想的（e.g., $\rho = 1$ ）でないため、相関の希薄化が生じる。相関係数の差を検討する際に尺度得点を用いると、たとえ差を見出したとしても、構成概念間の相関の差を検出したのか、各尺度の信頼性が一致しないことに由来する希薄化の程度差を検出したのか、またその両方の影響が混在したのか、不透明になる。そこで本研究では構造方程式モデリングにより潜在変数間の相関を検討することにより、相関の希薄化の程度差による問題を回避した。なお、構造方程式モデリングによる相関の希薄化の修正については、狩野（2002）が詳しく解説している。

附記

本研究は、日本教育心理学会第61回総会にて発表された内容に、新たな分析を施し、大幅な加筆を加えたものである。

謝辞

調査にご協力いただきましたA校、B校、C校の生徒および教職員の皆様にお礼申し上げます。

また、本研究の遂行に当たり高知大学大学院教育学専攻の宮崎亮介さんには多大なるお力添えを賜りました。ここにお礼を申し上げます。

本研究はJSPS科研費JP18H01017の助成を受けた。

文献

- 中央教育審議会（2015）：初等中等教育分科会（第100回）配布資料 資料1 教育課程企画特別部会 論点整理。
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/siryu/attach/1364306.htm（参照日2020.07.16）
- Eccles, J., & Wigfield, A. (2002): Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.
- 船越理沙, 田崎勝也, 潮村公弘（2013）：平均構造・多母集団同時分析を用いたセルフ・モニタリング（Self-Monitoring）尺度の文化的等価性の検討, *社会心理学研究*, 28, 3, 180–188.
- 濱保和治, 山崎敬人, 岡田大爾（2019）：理科学習の有用性を実感させるキャリア教育の実践的研究—理科学習における社会人講師活用の効果の考察を通して—, *理科教育研究*, 59, 3, 467–475.
- 平井洋子（2010）：測定の妥当性からみた尺度構成一得点

の解釈を保証できますか、吉田寿夫編著「心理学研究法の新しいかたち」, 21–49, 誠信書房。

市川伸一（1995）：学習と教育の心理学（現代心理学入門3）, 岩波書店。

伊田勝憲（2001）：課題価値評定尺度作成の試み, 名古屋大学大学院教育発達科学研究科紀要 心理発達科学, 48, 83–95.

科学技術振興機構（2019）：スーパーサイエンスハイスクール実施要項。 https://www.jst.go.jp/cpse/ssh/ssh/public/pdf/ssh_gaiyou.pdf（参照日2020.09.07）

科学技術振興機構（2020）：スーパーサイエンスハイスクール指定校一覧。 <https://www.jst.go.jp/cpse/ssh/school/list.html>（参照日2020.09.07）

鹿毛雅治（2013）：学習意欲の理論 動機づけの教育心理学, 金子書房。

狩野裕（2002）：構造方程式モデリングは、因子分析、分散分析、パス解析のすべてにとって代わるのか？, *行動計量学*, 29, 2, 138–159.

解良優基, 中谷素之（2014）：認知された課題価値の教授と生徒の課題価値評定, および学習行動との関連, *日本教育工学会論文誌*, 38, 1, 61–71.

解良優基, 中谷素之（2019）：課題価値のもつ概念的特徴の分析と近年の研究動向の概観, *南山大学紀要『アカデミア』人文・自然科学編*, 17, 95–116.

小林淑恵, 小野まどか, 荒木宏子（2015）：スーパーサイエンスハイスクール事業の俯瞰と効果の検証, 文部科学省 科学技術・学術政策研究所 DISCUSSION PAPER No. 117.

国立教育政策研究所（2017）：TIMSS2015 算数・数学教育/理科教育の国際比較, 明石書店。

国立教育政策研究所（2018）：平成30年度 全国学力・学習状況調査 報告書。 <https://www.nier.go.jp/18chousakekkahoukoku/report/question/>（参照日2019.07.16）

Konstabel, K., Lönnqvist, J. E., Walkowitz, G., Konstabel, K., & Verkasalo, M. (2012): The 'Short Five'(S5): Measuring personality traits using comprehensive single items. *European Journal of Personality*, 26, 1, 13–29.

熊谷龍一（2012）：統合的DIF検出方法の提案—“EasyDIF”の開発—, *心理学研究*, 83, 1, 35–43.

箕浦有希久, 成田健一（2013）：2項目自尊感情尺度の提案—評価と受容の2側面に注目して—, *人文論究*, 63, 1, 129–147.

文部科学省（2018）：高等学校学習指導要領（平成30年告示）解説 理科編 理数編。 http://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afiedfile/2019/03/28/1407073_06_1_1.pdf（参照日2020.07.16）

村上宣寛（2006）：心理尺度のつくり方, 北大路書房。

村山航（2012）：妥当性概念の歴史的変遷と心理測定学的観点からの考察, *教育心理学年報*, 51, 118–130.

Muthén, L. K., & Muthén, B. O. (2017): *Mplus user's guide*:

- Statistical analysis with latent variables, eighth edition.* Los Angeles, CA: Muthén & Muthén.
- Núñez-Peña, M. I., Guilera, G., & Suárez-Pellicioni, M. (2014): The single-item math anxiety scale: An alternative way of measuring mathematical anxiety. *Journal of Psychoeducational Assessment*, 32, 4, 306–317.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001): Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and social psychology bulletin*, 27, 2, 151–161.
- 田崎勝也 (2007)：文化的自己観は本当に「文化」を測っ

ているのか, 行動計量学, 34, 1, 79–89.

田崎勝也 (2008)：社会科学のための文化比較の方法等価性と DIF 分析, ナカニシヤ出版.

(受付日2020年7月29日；受理日2020年9月29日)

[問い合わせ先]

〒010-8502 秋田県秋田市手形学園町1-1

秋田大学教育文化学部

原田 勇希

e-mail: yukiharada@ed.akita-u.ac.jp
