

理科教育用メタ認知測定尺度の再考

—off-line メソッドの限界と今後に向けて—

原田 勇希^{1,2}
 久坂 哲也³
 草場 実²
 鈴木 誠⁴

【要 約】

学校教育においてメタ認知能力の育成は重要視されており、理科教育学においても個人のメタ認知能力を測定するための質問紙が開発されてきた。一方、近年の研究によると、質問紙法などの自己報告による測定方法（off-line メソッド）では、メタ認知の測定は難しいことが指摘されている。本研究ではこれまでに開発された理科教育用メタ認知測定尺度が、真にメタ認知を測定できているかを確かめるため、その収束的妥当性と弁別的妥当性を検討することを目的とした。研究の結果、メタ認知測定尺度と理科の学業成績（全国学力・学習状況調査）との間に実質有意な正の相関が確認できなかったことから、収束的妥当性は認められないと結論づけた。また自己愛傾向や社会的望ましき反応バイアスとの確かな正の相関が確認されたことから、弁別的妥当性は認められないと結論づけた。以上の結果をもとに、現在の理科教育学におけるメタ認知の測定方法の限界と今後の方策について考察した。

【キーワード】メタ認知, 測定, 妥当性, 質問紙法, ベイズ統計

1. 問題の所在

1.1 メタ認知の定義と意義

平成 29 年告示の学習指導要領における教育課程全体を通して育成を目指す資質・能力のうち、「学びに向かう力・人間性等」では「いわゆる『メタ認知』に関するもの」が挙げられており（文部科学省, 2016）、当然ながら理科においても育成が求められる。

メタ認知（metacognition）は「認知現象に対する知識と認知」（Flavell, 1979）と定義される。すなわち自分自身や他者の認知活動を一段上から認知することを指す概念である。メタ認知は、人・課題・方略についての知識を指す概念であるメタ認知的知識と、認知活動のモニタリングやコントロールを指す概念であるメタ認知的活動に大別でき、適切なメタ

認知的活動は適切なメタ認知的知識によって支えられる（e.g., 三宮, 2008）。

高いメタ認知能力を持つ学習者は、課題解決の最中において、適切に解決に向かっていくかを判断し、必要に応じて軌道修正ができる。そのためメタ認知は自己調整学習（self-regulated learning）の実現にとって欠かせない（三宮, 2010）。

教科教育においてメタ認知に介入する意義の根拠は、① 妥当に測定されたメタ認知能力が学業成績（認知的領域）と確かな正の相関を示すこと、② メタ認知への教育的介入が可能であること、の 2 点に集約できる。メタ認知と学業成績の間には確かな正の相関が確認されており（e.g., Dent & Koenka, 2016; Ohtani & Hisasaka, 2018）、この関連は知能による影響を考慮した上でも認められる（Ohtani & Hisasaka, 2018; van der Stel & Veenman, 2014; Veenman, Van Hout-Wolters & Afflerbach, 2006）。こうした事実は、メタ認知能力がまさに「学びに向かう力」として機能していることを示している。また、メタ認知への教育的介入による学業成績への効果は、発達段階や

¹ 日本学術振興会特別研究員 PD

² 高知大学教育学部

³ 岩手大学教育学部

⁴ 北海道大学大学院理学院

教科による効果量の違いもみられるが、正の影響が認められる (Dignath & Büttner, 2008)。そのため、理科に限らず教科教育では学習者のメタ認知に積極的に介入し、これを伸長することが重要である。

メタ認知能力の領域固有性について van der Stel & Veenman (2014) は、発達 (12 歳→15 歳) に伴って領域一般的な能力として統合される傾向があるが、発達の初期では領域固有性が認められることを報告している。また、理科でのメタ認知の重要性について Veenman (2012) は、理科学習の特徴である科学的な探究過程が複数の認知プロセスの組み合わせによって形成されていることから、適切なモニタリングとコントロールが重要であると述べている。これと類似した指摘は、湯澤 (2008) によってもなされている。そのため、理科教育学でも子どものメタ認知能力の向上を目的とした研究が盛んに行われてきた (レビューとして、久坂, 2016)。理科教育学研究では、将来的な領域一般的能力への転移を見据えながら、“理科学習におけるメタ認知”の能力を育成することが重要であり、これを適切に測定、評価する必要がある。

1.2 メタ認知の測定 (評価) に伴う問題

理科教育学においてメタ認知の働きや役割に目を向けることは重要であるが、他の心的概念と比して構成概念が曖昧であり、しばしば測定の困難さが問題になる (久坂, 2016)。例えば、質問紙によるメタ認知の測定では、回答者に“自分のメタ認知の働きをメタ認知”することを要求する。そのため正しい測定の条件は回答者が高い“メタ・メタ認知能力”を持っていることであると考えられることから、メタ認知の測定が一筋縄でいかないことは想像するに難くない。

メタ認知の測定方法は、課題遂行中の発話や行動、および生理指標などを記録し分析する on-line メソッドと、課題遂行の事前または事後段階において日常の学習場面あるいは特定の課題遂行場面を想起させて回答させる off-line メソッドに大別される (久坂, 2016; Veenman et al., 2006)。代表的な on-line メソッドには発話思考法や行動観察法が挙げられる。また代表的な off-line メソッドには質問紙法やインタビューなどが挙げられる。

両測定方法とも、同じく“メタ認知”を測定しているはずであるが、量的分析 (相関分析など) を行ったときの両者の振る舞いは以下のように大きく異なる (e.g., Bannert & Mengelkamp, 2008; 久坂, 2016; Veenman, 2005)。第 1 に、on-line メソッドと

off-line メソッドの間の相関は弱い。第 2 に、on-line メソッドに該当する複数の方法間の相関は強いのに対し、off-line メソッドに該当する複数の方法間の相関は弱い。第 3 に、on-line メソッドと学業成績の相関は強いのに対し、off-line メソッドと学業成績の相関は弱い。こうした知見が蓄積された現在では、off-line メソッドはメタ認知を正確に捉えていない可能性が高いと考えられている (Veenman, Bavelaar, De Wolf & Van Haaren, 2014)。国内の研究においても、松浦・角屋・岡田・檜山 (2002) による先駆的な研究では、作成したメタ認知尺度と電磁気領域の知識・理解との間に弱い相関 ($r=.273$) しかみられず、面接法 (on-line メソッドに該当する方法) によって測定したメタ認知との間に有意な関連がみられないことから、質問紙法によってメタ認知を測定することは困難であると考察している。

1.3 これまでの理科教育学研究における測定

これまでに理科教育学におけるメタ認知を扱った研究を系統的にレビューした久坂 (2016) によると、質問紙法による測定が 67.9% と最も多く、off-line メソッドに偏重している現実がある。『理科教育学研究』や関連領域の学術誌では測定尺度を作成した論文が数多く掲載されている (e.g., 木下・松浦・角屋, 2005; 草場・足達・鈴木, 2017; 松浦・木下, 2008; 宮本・木下・網本, 2015; 鈴木, 1997)。

前述したように、off-line メソッドはメタ認知を正しく捉えられないと考えられている。それでは、これまでに作成された理科教育用メタ認知測定尺度 (以下、メタ認知尺度) は何を測定していたのだろうか。通常、心理尺度が作成される場合には信頼性と妥当性が検討される。とりわけ、妥当性とは当該尺度が測定したい構成概念を捉えているかを指す概念であり、off-line メソッドの問題の所在である。メタ認知尺度が開発された際に十分に妥当性が検討されているならば、そのプロセスを参照することで、真にメタ認知を測定できているかを確かめられると期待される。そこで以下では、理科全般を測定対象とするメタ認知尺度を開発した代表的な研究の妥当性検討プロセスを概観する。

鈴木 (1997) は理科教育学の立場より“認知的方略のメタ認知測定尺度 (Metacognition Scale of Cognitive Strategies; MSCS)”を開発している。本尺度の開発時には妥当性の検討を行っていないが、MSCS の短縮版を作成する際に、「基準関連妥当性を調べるために」他のメタ認知測定尺度と正の相関 ($r=.39, .47$) があることを根拠に「妥当性があるものと考え

られる」としている(鈴木, 1999)。

木下ら(2005)では、「質問項目の妥当性を検討するため、主成分分析を」を行い、抽出された主成分が作成時に意図した構造と一致したことを根拠に「作成した質問項目は妥当性があると判断」している。また、項目の作成にあたり「小学校教師7名、中学校理科担当教師3名、理科教育学を専攻する大学院生3名で検討」と記述されていることから、内容的妥当性の検討も行われたと判断できる。また、この尺度を発展させた松浦・木下(2008)は、木下ら(2005)と同様の手続きで項目を作成、追加し、「項目の得点と全体の得点」のシリアル相関が高いことを持って、「妥当性が得られたと判断」している。

草場ら(2017)では、メタ認知の理論から導かれるように、「メタ認知的知識」、「モニタリング」、「コントロール」の3つの潜在変数と、「メタ認知」を高次因子として設定した因子構造の適合度が十分であることを根拠に「妥当性が担保されている」としている。また、「理科教育学を専門とする大学教員1名、高等学校の現職理科教員1名及び理科教育学を専攻する大学院生2名」が項目を作成したと記述されていることから、内容的妥当性の検討も行われたと判断できる。

以上のように、これらのメタ認知尺度は作成時点に様々な妥当性検討の過程を経ている。一方、それぞれの研究で確認されているのは妥当性概念の一部に過ぎないことも事実である。メタ認知尺度の構成概念妥当性を慎重に検討するには、理論的に関連が想定される理科の学業成績との正の相関や(i.e. 収束的妥当性)、理論的に関連しないと考えられる他変数との無相関(i.e. 弁別的妥当性)を確認する必要があるだろう。しかし、これまでの研究においてメタ認知尺度の収束的妥当性と弁別的妥当性は検討されていない。そのため現状では、off-line メソッドであるメタ認知尺度が、近年の研究で指摘されているようにメタ認知を正しく捉えられていないのか、それともこうした批判は当たらないものであるのか、不透明である。

2. 本研究の目的と作業仮説

こうした背景を受け、本研究では心理学的な方法論に依拠し、これまでの理科教育学領域で開発されてきたメタ認知尺度の収束的妥当性と弁別的妥当性の検討を目的とする。

本研究では収束的妥当性について、メタ認知尺度と学業成績の相関についてメタ分析した先行研究(Ohtani & Hisasaka, 2018)に倣い、認知的領域に注

目した学業成績の指標を扱う。もしメタ認知尺度に収束的妥当性があるならば、理科の学業成績と確かな正の相関が確認できると考えられる。また、もしメタ認知尺度に弁別的妥当性があるならば、メタ認知と理論的に関連しない心理変数と無相関であると考えられる。

3. 研究1：理科教育用メタ認知測定尺度と理科の学業成績、および自尊心、自己愛傾向、社会的望ましき反応バイアスとの相関分析

3.1 序論

本研究では、メタ認知尺度の収束的妥当性と弁別的妥当性を相関分析によって検討する。

収束的妥当性の検討に使用する理科の学業成績の指標として、本研究では平成30年度に実施された全国学力・学習状況調査(国立教育政策研究所, 2018a)を使用する。以降の議論では、当調査問題の得点には十分に“真の理科の学力”が反映されているものとみなす¹⁻³⁾。その根拠は、当調査問題は「主として『知識』に関する問題」と「主として『活用』に関する問題」を含み、評価の観点や出題領域も幅広く、また悉皆調査であるため数多くある学業成績指標の中で最も標準化された指標であると考えられるためである。

弁別的妥当性の検討に使用する変数として、あえて理科とは関連のない心理変数を取り上げる。具体的には、自尊心、自己愛傾向、および社会的望ましき反応バイアスを測定する。

自尊心は、自己に対する肯定的な感覚を指す概念であり、自尊心と学業成績の間にはほとんど関連がないことが明らかになっている(Baumeister, Campbell, Krueger & Vohs, 2003)。領域一般的な概念である自尊心と理科のメタ認知との間に関連があるとすれば、動機づけを媒介する経路が考えられる。しかし、その効果は個人によって異なる自己価値の随伴性(e.g., Crocker & Wolfe, 2001)による影響を受けると考えられるため(e.g., 大谷・中谷, 2011, 2013)、全体としてみると関連は小さいと予想される。

自己愛傾向は、自身に対する誇大な感覚、賞賛欲求、共感性の欠如といった感覚を抱く傾向を指す概念である(小塩, 2012)。自尊心と自己愛傾向は、自己に対する肯定的感覚という点で類似しており、正の相関がある(岡田, 2009)。自己愛傾向に固有な特徴のひとつに自分自身を過剰に高く評価することが挙げられる(e.g., American Psychiatric Association, 2013 高橋・大野監訳, 2014; 小塩, 2012)。実際、自己愛傾向が高い個人ほど自分の知的能力を過剰に

高く評価する傾向があり、この現象は自尊心よりも自己愛傾向で顕著である (Gabriel, Critelli & Ee, 1994; 中村, 2018)。

以上のことより、もしメタ認知尺度と自己愛傾向との間に正の相関があり、かつその相関がメタ認知尺度と自尊心との相関よりも強いのであれば、メタ認知尺度に対する回答には誇大な自己感覚が反映されていると推測できる。すなわち、メタ認知尺度の弁別的妥当性を疑わざるをえない。

社会的望ましき反応バイアスとは、質問紙などの回答に際し、意識的あるいは無意識的に社会的に望ましいとされる回答を選択する傾向を指す。社会的望ましき反応バイアスはしばしば測定尺度の妥当性を侵害することが知られている。もしメタ認知尺度と社会的望ましき反応バイアスとの間に正の相関があるならば、正しくメタ認知を測定できていないと考えられる。

3.2 方法

3.2.1 研究対象者

A 県の X 中学校に通う 3 年生 ($n = 131$) を対象とした。

3.2.2 測定変数 (附録)

メタ認知尺度 本研究では以下 3 つのメタ認知尺度を検討した。1 つ目は、鈴木 (1997) によって作成された MSCS の短縮版 (ver. 2; 鈴木, 1999) である。本尺度にはメタ認知ではなく自己効力感に該当するような質問項目が含まれていることから (e.g., 理科の授業の内容は、私にはかんたんです)、現在の知見に基づきメタ認知のモニタリングに焦点を当てていると考えられる 4 項目を使用した。回答は“1. 全く当てはまらない”から“5. よく当てはまる”の 5 件法で求めた (以下、同様)。

2 つ目は、松浦・木下 (2008) によって作成された尺度である。この研究では、木下ら (2005) による尺度に同様の手順で新たな項目を追加し、項目反応理論によって等価な 2 つの尺度を作成している。本研究ではその一つを構成する 16 項目全てを使用した。

3 つ目は、草場ら (2017) によって作成された尺度である。この尺度はメタ認知の理論に基づき、メタ認知的知識、モニタリング、コントロールの 3 つを測定範囲としている。本研究では 12 項目全てを使用した。

理科の学業成績 平成 30 年度に実施された全国学力・学習状況調査 (国立教育政策研究所, 2018a) の

正答率を指標とした。

自尊心 Rosenberg (1965) による尺度から、加藤・太田・松下・三井 (2018) で使用されている 4 項目を使用した。

自己愛傾向 中山・中谷 (2006) の評価過敏性-誇大性自己愛尺度より、誇大性の下位尺度から 4 項目を選択して使用した。

社会的望ましき反応バイアス 谷 (2008) によるバランス型社会的望ましき反応尺度日本語版 (Japanese version of Balanced Inventory of Desirable Responding; BIDR-J) の自己欺瞞の下位尺度より 5 項目を選択して使用した。自己欺瞞とは、本当に自分の自己像であると信じて無意識的に社会的に望ましく回答する反応である。

3.2.3 調査手続き

質問紙調査は 2018 年の 10 月に理科授業内で行われた。調査協力が得られた質問紙を有効回答として分析に使用した。また成績には関係しないことが教示された。

本調査への協力に際し協力校への十分な説明を行い、職員会議にて同意の決定を得た。全国学力・学習状況調査の結果は協力校内で教諭の監督のもと転記・紐付けを行い、個人情報を持ち出しは行わなかった。また、調査結果の研究利用と論文文化について文部科学省に助言を求め、指示に従った。

3.2.4 統計的分析

本研究では帰無仮説検定を使用しない。その理由は、本研究ではメタ認知尺度の妥当性を検討するが、学業成績との無相関 ($r = 0$) が棄却されることだけでは収束的妥当性があると判断できないためである。

メタ分析によると、on-line メソッドでは学業成績と $r = .53$ [.45, .61], off-line メソッドでは $r = .23$ [.20, .26] の相関がある (Ohtani & Hisasaka, 2018)。つまり、off-line メソッドも学業成績と正の相関があるが、問題はその相関の弱さにある。頻繁に使用される帰無仮説検定の使用方法では、無相関でないことは検討できるが、相関係数が .20 以上である確率 (以下、 P ($r > .20$) と表記) など、パラメータが任意の値以上である確率を求められない。また、 p 値の性質よりサンプルサイズを大きくするほど小さな効果量でも有意になるため、検定結果 ($p < .05$) を収束的妥当性の判断基準とすると、 $n = 146$ を用意することで、母相関係数が .23 であれば 80% の確率で収束的妥当性があると結論づけられてしまう (G*Power3 (Faul, Erdfelder, Lang & Buchner, 2007)

による検定力分析に基づく)。

そこで、本研究ではベイズ推定法によって求めた相関係数の事後分布を吟味し、メタ認知尺度の妥当性を議論する。具体的な基準を以下に述べる。まず従来の帰無仮説検定と類似した意思決定の方法として、95%両側確信区間(95% Credible Interval; 以下、95%CI)が0を跨いでいないことをもって相関係数の正負を判断する。弁別的妥当性については相関係数の目安がないことから、この方法で判断する。

収束的妥当性について、強い意味ではon-lineメソッドと同程度の正の相関が検出されることが望ましい。しかし、メタ認知尺度が簡便法として用いられることを考慮し、on-lineメソッドよりも弱い相関であっても認めることとした。具体的にはOhtani & Hisasaka (2018) が示した相関係数の下限値を参考にし、 $r = .20$ を最低基準とする。そのため収束的妥当性の有無の判断は $P(r > .20)$ が十分に大きいことを確認して行う。また収束的妥当性の程度の判断は、得られた相関係数の点推定値や $P(r > .30)$ などをもとに総合的に行う。

統計分析や図の出力にはR (ver 3.5.3) とパッケージ rstan (ver 2.18.2), loo (ver 2.1.0), bayesplot (ver 1.6.0) を用いた。反復回数 80,000 のチェーンを4つ発生させ、反復の前半半分をバーンイン期間とした(マルコフ連鎖モンテカルロ法)。モデルは多くの先行研究と同様、各測定変数が正規分布(i.e. 2変量正規分布)に従うものとし、Pearsonの積率相関係数の事後分布を求めた。事前分布にはデフォルトの無情報分布を採用した。収束判断は $\hat{R} < 1.01$ とし、全てのパラメータがこれを満たすことを確認した。

3.3 結果と考察

メタ認知尺度の因子構造を確認した。鈴木(1997)による尺度の固有値は2.80, 0.53, 0.42…と減衰しており、第1因子の寄与率は69.95%であった。松浦・木下(2008)による尺度の固有値は11.08, 1.10, 0.84…と減衰しており、第1因子の寄与率は69.26%であった。草場ら(2017)による尺度の固有値は8.32, 0.67, 0.57…と減衰しており、第1因子の寄与率は69.35%であった。これらの結果より、3つの尺度とも1因子構造であると判断し、項目平均値を尺度得点とした。

表1に各測定変数の記述統計量、相関係数と95%CIを示した。内的整合性は自尊心がやや低い傾向であったものの、おおむね十分な値を示した。

メタ認知尺度の収束的妥当性を検討するために、学業成績との間の相関係数を確認したところ、どの尺度も95%CIに0を含んでおり、確かな正の相関は確認できなかった。それぞれの尺度について $P(r > .20)$ を求めたところ、鈴木(1997)による尺度は3.66%、松浦・木下(2008)による尺度は3.22%、草場ら(2017)による尺度は5.48%であった。この結果は、メタ認知尺度は理科の学業成績と.20以下の相関しか持たない確率が高いことを意味しており、メタ認知尺度の収束的妥当性に疑問を投げかけるものである。

メタ認知尺度の弁別的妥当性を検討するために、自尊心、自己愛傾向、社会的望ましき反応バイアスとの間の相関係数を確認したところ、確かな正の相関が認められた。また、どの尺度も自尊心よりも自己愛傾向と強く相関する傾向があった。そこで、メ

表1 各測定変数の記述統計量と内的整合性、相関分析の結果

測定変数	n	記述統計量			相関分析						
		Mean	(SD)	ω	1	2	3	4	5	6	
1 理科の学業成績	131	0.00	(1.00)								
2 鈴木(1997)	126	3.72	(0.92)	.859	.024						
					[-.172, .216]						
3 松浦・木下(2008)	124	3.86	(0.83)	.970	.017	.802					
					[-.177, .211]	[.730, .862]					
4 草場・足達・鈴木(2017)	125	3.73	(0.90)	.959	.041	.815	.909				
					[-.153, .234]	[.743, .870]	[.871, .937]				
5 自尊心	125	2.81	(0.80)	.660	-.083	.192	.273	.242			
					[-.274, .110]	[.012, .366]	[.095, .439]	[.061, .414]			
6 自己愛傾向	126	2.62	(1.11)	.920	-.117	.298	.370	.358	.631		
					[-.306, .075]	[.122, .462]	[.198, .521]	[.187, .510]	[.510, .734]		
7 社会的望ましき 反応バイアス	121	2.83	(0.97)	.867	-.158	.375	.351	.339	.461	.695	
					[-.345, .037]	[.204, .530]	[.177, .508]	[.165, .498]	[.301, .599]	[.589, .783]	

注) 相関係数の点推定値として、事後分布のEAP(Expected A Posteriori)を求めた。[]内に95%CIを記載した。理科の学業成績は協力校の匿名性を考慮し、標準化して分析した。

メタ認知尺度と自尊心、自己愛傾向との関係について、どちらに固有な成分とより密接に関連するかを検討するため、片方を制御変数とする偏相関係数 (pr) を求めた。自己愛傾向を制御変数としたメタ認知尺度と自尊心との偏相関係数は、鈴木 (1997) が $pr_{EAP} = .005 [-.176, .185]$ 、松浦・木下 (2008) が $pr_{EAP} = .055 [-.124, .236]$ 、草場ら (2017) が $pr_{EAP} = .023 [-.156, .201]$ であり、いずれも確かな偏相関は認められなかった。一方、自尊心を制御変数としたメタ認知尺度と自己愛傾向との偏相関係数は、鈴木 (1997) が $pr_{EAP} = .238 [.061, .405]$ 、松浦・木下 (2008) が $pr_{EAP} = .265 [.090, .428]$ 、草場ら (2017) が $pr_{EAP} = .271 [.098, .432]$ であり、いずれも確かな正の偏相関が認められた。この偏相関分析の結果は、自己愛傾向に固有な成分とメタ認知尺度が確かな相関を持つことを示しており、メタ認知尺度の弁別的妥当性に疑問を投げかけるものである。

以上より、これまで理科教育学領域で開発されてきたメタ認知尺度は収束的・弁別的妥当性に乏しく、理科学習におけるメタ認知を正しく捉えられていない可能性が高いといえる。

しかしながら、本研究で示した学業成績との無相関性について、以下のように批判できる。メタ認知と学力の関連を検討する場合、難易度などの課題の性質に注意する必要がある (e.g., 松浦ら, 2002)。なぜなら、メタ認知を要求しない単純な問題であれば、即座に認知レベルで解に至るため、メタ認知を働かせる必要がないからである。例えば、全国学力・学習状況調査の問題は4つの評価観点から構成されるが、「自然事象についての知識・理解」や「観察・実験の技能」の問題では、学習時に記録、テスト時に検索し再生するだけの記憶課題の手続きで解けた可能性がある。科学的思考にはメタ認知が必要不可欠であることを鑑みると (e.g., Veenman, 2012; 湯澤, 2008)、メタ認知が寄与する学業成績の観点は「科学的な思考・表現」であると考えられる。他の観点が混在した本研究ではそれがノイズとなり、本来あるはずの正の相関が検出できなかった可能性がある。

4. 研究2：理科の学業成績の各評価観点と理科教育用メタ認知測定尺度の相関分析

4.1 序論

研究1では、学業成績として全体の正答率のみを取得したため、これ以上の分析は不可能であり、前述の批判に回答できなかった。そこで、研究2では評価観点ごとに相関係数を求め、メタ認知尺度の収束的妥当性について詳しく検討する。

4.2 方法

4.2.1 研究対象者

研究1と同年度に中学校3年生を対象とした調査を行うと、一度問題を解いているため、結果に歪みが生じる可能性がある。そこで研究2では中学校2年生を対象に調査を行うこととし、A県のX中学校に通う2年生 ($n=111$) を対象とした。

4.2.2 測定変数

メタ認知尺度 研究1と同様に、鈴木 (1999)、松浦・木下 (2008)、草場ら (2017) による項目を使用した。すべての項目が研究1と同一であった。

理科の学業成績 平成30年度に実施された全国学力・学習状況調査の問題 (国立教育政策研究所, 2018a) を実施した。問題番号6の学習内容が未習であったことから、調査問題から除外した。各評価の観点の得点は、国立教育政策研究所 (2018b) による「問題別集計結果」に示されている問題の合計正答数を求めた。

4.2.3 調査手続き

質問紙調査は2018年の2月に理科授業内で行われた。調査協力が得られた質問紙を有効回答として分析に使用した。また成績には関係しないことが教示された。

4.3 結果と考察

表2に各測定変数の記述統計量と内的整合性を示した。「自然事象についての知識・理解」の内的整合性が許容できない水準であった。当評価観点内には多様な出題領域や問題形式が含まれている。帯域幅

表2 各測定変数の記述統計量と内的整合性

測定変数	n	記述統計量		ω
		Mean	(SD)	
理科の学業成績				
1 総正答数	111	15.39	(3.51)	.717
2 自然事象への関心・意欲・態度	111	0.33	(0.47)	—
3 科学的な思考・表現	111	9.47	(2.66)	.678
4 観察・実験の技能	111	1.32	(0.62)	—
5 自然事象についての知識・理解	111	4.59	(1.06)	.304
メタ認知尺度				
6 鈴木 (1997)	106	3.29	(0.67)	.750
7 松浦・木下 (2008)	106	3.40	(0.68)	.937
8 草場・足達・鈴木 (2017)	106	3.32	(0.68)	.908

注) 自然事象への関心・意欲・態度は1問、観察・実験の技能は2問であったため、 ω 係数は算出しなかった。

と忠実度のジレンマを考慮すると、出題の多様性によって「自然事象についての知識・理解」という抽象度の高い構成概念の測定の妥当性を高めていることが考えられた。これを考慮し、以降では学業成績の内的整合性については不問とした。

まず、メタ認知尺度の収束的妥当性と研究1の結果の再現性を検討するために、学業成績の総正答数との間の相関係数を確認した(表3)。どの尺度も95%CIに0を含んでおり、確かな正の相関は確認できなかった。それぞれの尺度について $P(r > .20)$ を求めたところ、鈴木(1997)による尺度は42.49%、松浦・木下(2008)による尺度は27.72%、草場ら(2017)による尺度は31.34%であった。メタ認知尺度と学業成績全体との間の相関係数は研究1よりやや大きい傾向であったが、それでも $r_{EAP} < .20$ であり、 $P(r > .20)$ が50%以下であることから、研究1とおおむね同様の結果が再現されたと判断した。

次に、メタ認知尺度と学業成績の観点ごとの正答数の相関係数を確認した。3つのメタ認知尺度とも、すべての観点との間に確かな正の相関は確認できず、相関係数のEAP推定値は小さかった($r_{EAP} < .106$)。また、評価の観点による相関係数の顕著な差異はみられなかった。

結果を参照すると、メタ認知尺度が理科の「科学的な思考・表現」に関する学業成績と正の相関があるとはいえない。すなわち、全国学力・学習状況調査の正答率や総正答数と確かな相関がない背景として考えられた、「自然事象についての知識・理解」や「観察・実験の技能」などのメタ認知を働かせる必要のない観点が混入したという仮説は成立しない。よって、研究2の結果も研究1の結果と同様に、メタ認知尺度の収束的妥当性に疑問を投げかけるものである。

5. 研究3：研究1, 2の結果を統合した事後分布の定義による本研究の結論の導出

5.1 序論

研究1と研究2では、一貫してメタ認知尺度と理科の学業成績(研究1：正答率, 研究2：総正答数)

との間にある確かな正の相関を検出できなかった。しかし、いずれも $r_{EAP} < .20$ ではあるものの、研究1と比較して研究2の方が、得られた相関係数がやや大きい傾向であった。研究1と2では学年(研究1：中学校3年生, 研究2：中学校2年生)、問題(研究1：全問題, 研究2：問題番号6を除いた)で条件が異なっていたことを考慮すると、研究1と2で母相関係数が同一でない可能性がある。

そこで、研究1と研究2において、母相関係数が同一であるモデル(モデル1： $\rho_1 = \rho_2$)と、異なるモデル(モデル2： $\rho_1 \neq \rho_2$)の情報量規準の比較を行う。モデル2が採択されるのであれば、メタ認知尺度と理科の学業成績との相関は、何らかの要因によって調整されることを示唆する。反対に、モデル1が採択されるのであれば、メタ認知尺度と理科の学業成績との相関は、本研究の条件下においては同一であることを示唆する。

またモデル1が採択された場合には、収束的妥当性についての結論を下すため、研究1と研究2の結果を統合し、採用したモデルの下における相関係数の事後分布を定義する。

5.2 結果と考察

表4にモデル1とモデル2の各情報量規準を示した。WAICを参照すると、鈴木(1997)はモデル2、草場ら(2017)はモデル1が支持される傾向であった。しかし、WAICの標準誤差(se)と比較して差異は非常に小さかった。DICは3つの尺度とも僅かにモデル2を、BICは3つの尺度とも僅かにモデル1を支持していた。

この結果より、モデルの優劣についての明確な証拠は得られなかったと判断した。以降では、モデルの儉約性と、現時点では相関係数の値を調整するメカニズムについて有力な仮説がないことを考慮し、モデル1を採択することとした。

モデル1のもと、研究1と研究2の結果を統合した相関係数の事後分布とトレースプロットを図1に、EAP推定値と95%CIおよび相関係数が0, .10, .20, .30以上である確率を表5に示した。3つのメタ認知尺度とも、収束的妥当性の最低基準として設定した.20

表3 メタ認知尺度と理科の学業成績との間の相関係数

メタ認知尺度	総正答数	各評価の観点			
		関心・意欲・態度	思考・表現	技能	知識・理解
鈴木(1997)	.179 [-.022, .368]	.083 [-.129, .294]	.070 [-.152, .298]	.100 [-.105, .298]	.056 [-.160, .265]
松浦・木下(2008)	.139 [-.062, .332]	.058 [-.149, .270]	.105 [-.121, .320]	.101 [-.104, .300]	-.035 [-.243, .179]
草場・足達・鈴木(2017)	.149 [-.051, .341]	.006 [-.200, .217]	.092 [-.130, .313]	.106 [-.096, .305]	.024 [-.189, .238]

表 4 各モデルの情報量規準

		WAIC	(se)	DIC	BIC
鈴木 (1997)	モデル 1	1318.548	(41.319)	1315.420	1345.408
	モデル 2	1317.704	(41.023)	1314.672	1349.410
松浦・木下 (2008)	モデル 1	1309.177	(40.666)	1306.114	1336.356
	モデル 2	1309.177	(41.143)	1306.021	1340.745
草場・足達・鈴木 (2017)	モデル 1	1312.236	(40.184)	1309.565	1339.578
	モデル 2	1312.493	(40.582)	1309.550	1344.284

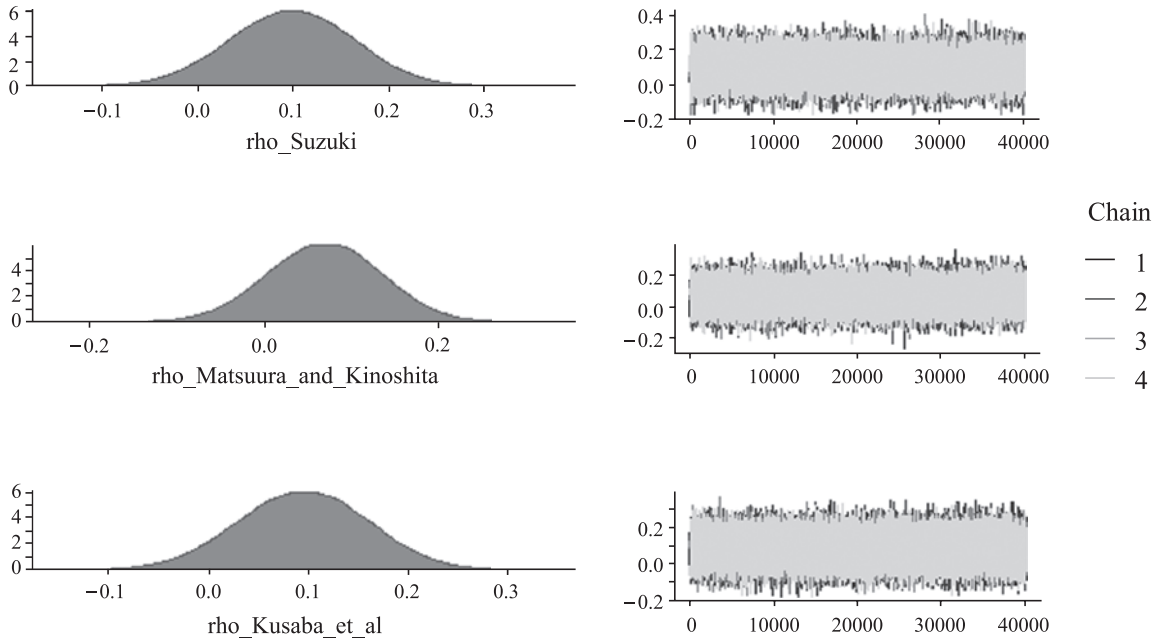


図 1 各メタ認知尺度と理科の学業成績との間の相関係数の事後分布（左）とトレースプロット（右）

注) 事後分布はバーンイン期間を除いた 160000 個の MCMC サンプルによって近似されたものである。上から鈴木 (1997), 松浦・木下 (2008), 草場ら (2017) である。事後分布の縦軸は確率密度, 横軸は相関係数を示している。図示したトレースプロットにはバーンイン期間は含まれていない。

表 5 メタ認知尺度と理科の学業成績との間の相関係数

	相関係数			確率 (%)			
	r_{EAP}	(post.sd)	95%CI	$P (r > .00)$	$P (r > .10)$	$P (r > .20)$	$P (r > .30)$
鈴木 (1997)	.098	(0.065)	[-.031, .225]	93.240	49.215	5.848	0.074
松浦・木下 (2008)	.070	(0.066)	[-.060, .197]	85.525	32.570	2.278	0.025
草場・足達・鈴木 (2017)	.093	(0.066)	[-.037, .220]	92.108	46.154	5.016	0.057

以上の相関がある確率はおよそ 5% 程度であり, .30 以上の相関がある確率はほぼ 0% であった。

6. 総合考察

6.1 理科教育用メタ認知測定尺度の妥当性

本研究の目的は, これまでの理科教育学領域で開発されてきたメタ認知尺度の収束的妥当性と弁別的

妥当性を検討することであった。それぞれの妥当性について以下に考察する。

メタ認知尺度の収束的妥当性について, 研究 3 で統合された学業成績との相関係数の事後分布を参照すると, 最低基準として設定した .20 以上である確率はおよそ 5% 程度であった。off-line メソッドによるメタ認知の指標と学業成績の相関をメタ分析

した Ohtani & Hisasaka (2018) は、相関係数を .23 [.20, .26] と見積もっている。これを考慮すると、国内の理科教育学領域で開発されたメタ認知尺度は Ohtani & Hisasaka (2018) が示した相関係数の下限値よりも低いことを示唆しており、当然ながら on-line メソッドによる相関係数 ($r=.53$ [.45, .61]) には遠く及ばない。また研究 2 で評価の観点別に相関係数を求めたところ、どの観点でも確かな正の相関は確認できなかった。以上の結果より、本研究ではメタ認知尺度の収束的妥当性は認められない可能性が高いと結論づける。

メタ認知尺度の弁別的妥当性について、研究 1 の結果より、自己愛傾向、社会的望ましき反応バイアスとの間に確かな正の相関があることが明らかになった。どちらの変数とも理科学習との関連は想定しづらく、理論的観点からメタ認知を促進する要因であるとは考えられない。自己愛傾向が高い個人ほど自らの知的能力を過大評価する傾向があることから (e.g., Gabriel et al., 1994), メタ認知尺度には誇大な自己感覚が反映されていると考えられる。同時に、本研究では社会的望ましき反応バイアスの指標として自己欺瞞を使用したことから、「本当に自分の自己像であると信じて無意識的に」(谷, 2008), 自分は理科の学習でメタ認知ができていと回答する者がいると考えられる。以上の結果より、本研究ではメタ認知尺度の弁別的妥当性は認められない可能性が高いと結論づける。

6.2 理科教育学研究の今後の方向性

新学習指導要領で育成を目指す資質・能力のうち、「学びに向かう力・人間性等」では「いわゆる『メタ認知』に関するもの」が挙げられており(文部科学省, 2016), また科学的思考力にはメタ認知が必要不可欠であることから (Veenman, 2012; 湯澤, 2008), たとえ測定が難しいとしても、理科教育学としてメタ認知の研究を進めていく必要がある。本研究の結果と久坂 (2016) によるレビューより、将来の理科教育学におけるメタ認知研究の方向性として以下 2 つが考えられる。

1 つ目は、理科学習におけるメタ認知を on-line メソッドで捉えていく方向である。on-line メソッドはより妥当にメタ認知を捉えることができるため、研究者にとって有用である。例えば、松浦ら (2002) では面接法によって測定したメタ認知(やり直し、予想・計画の変更、まとめ)が実験の成功/失敗と有意に関連することを示しており、この研究は on-line メソッドでのメタ認知の測定に成功した例と

いえよう。しかしながら on-line メソッドは、発話プロトコルや行動データを収集し、コーディングによって量的な指標に変換する必要がある、大きなコストがかかる。そのため理科教師にとって必ずしも有用とはいえない。off-line メソッドの精度が低いとしても、許容できる水準でメタ認知を反映しているのであれば簡便法として有用であろうが、本研究では最低基準として設定した収束的妥当性も確認できなかった。そのため、教師が off-line メソッドによって子どものメタ認知の働きを把握した場合、誤った能力評価に基づいた授業実践や教育的介入がなされる危険性があることから、現時点において off-line メソッドによる診断的評価、および形成的評価は推奨できないと考える。

2 つ目は on-line, off-line を問わず、より簡便かつ妥当にメタ認知を測定できる手法開発への方向である。現在の研究状況では、清野・及川・古本・久坂 (2018) が簡便な方法であるふきだし法を用いて on-line でのメタ認知の測定を試みているが、十分な精度でメタ認知を測定できているとは言い難く、未だ改善の余地がある。off-line メソッドでメタ認知を測定することに成功した数少ない例である Schellings, van Hout-Wolters, Veenman & Meijer (2013) による研究では、測定対象とするメタ認知の領域固有性を課題レベルまで十分に絞り、かつ実際に得られた発話プロトコルを基に質問項目を作成することで、on-line メソッドによる指標と高い相関 ($r=.63$) が得られることを報告している。この試みを参考に、課題固有のメタ認知尺度を開発していくべきだろう。

2 つ目の方向性に進んだ結果、妥当性の高いメタ認知尺度が完成したならば、それは理科教育学におけるメタ認知研究のブレークスルーといえるだろう。しかし、研究領域全体がこの方向を目指すとその間メタ認知の働きや役割を明らかにする研究が滞ることになる。そのため、研究領域全体としては両方向性を両立することが重要だろう。

6.3 本研究の限界

本研究ではメタ認知尺度について、収束的妥当性、弁別的妥当性とも認められないと結論づけた。しかし、収束的妥当性についての本研究の結論は、全国学力・学習状況調査の得点には“真の理科の学力”が十分に反映されているという仮定に支えられている。もし、全国学力・学習状況調査の得点が“真の理科の学力”と無相関であるなら(全国学力・学習状況調査が“真の理科の学力”を全く反映しないテストであるなら)、メタ認知尺度と“真の理科の学

力”の母相関係数がいかなる値であろうと、本研究で示した結果が導かれる²⁾。あるいは、“真の理科の学力”が多次元的であり、メタ認知尺度が全国学力・学習状況調査で測定した学力以外の側面とは十分に高い相関を示す可能性は否定できない³⁾。

本研究は現実的に実施可能な方法で行われ、全国の中学生を対象とした無作為抽出が徹底できたわけではない。これを受け、本研究で推定した相関係数の信頼性について考察する。本研究の対象校は1校であり、同一学年は1人の理科教師が担当していたため、教師要因による影響はそれぞれの学年内では一様とみなせる。そのため、メタ認知尺度や学業成績の平均値や標準偏差の推定値に教師要因や学校風土による影響が混入した可能性は否定できない。しかし、現時点においてメタ認知尺度と学業成績の個人差の連動具合（＝相関関係の強さ）の教師要因による調整効果（e.g., マルチレベル分析におけるクロス水準の交互作用など）は想定しづらい。実際、研究3では相関係数の調整効果に関する証拠は見出されなかった上、現在のところこれを報告した先行研究は見当たらない。しかし、本研究で検討しなかった何らかの調整変数が存在しないとは断定できないため、より強い確信度の結論を導くためには多くの追試とメタ分析が必要である。

いくつかの研究では教育的介入の前後でメタ認知尺度の得点が向上することを報告しているが（e.g., 木下, 2010; 草場・湯澤・角屋・森, 2010; 草場・角屋・森, 2012; 宮本ら, 2015）、この差分が何を意味するかについて慎重に再検討する必要がある。メタ認知尺度に混入していると考えられる自己愛傾向が理科の授業で変動するとは考えにくく、さらにその効果が「統制群」よりも「処置群」や「実験群」で顕著である可能性は低いだろう。そのため、メタ認知能力そのものを測定することは困難であったとしても、この差分には“獲得した理科のメタ認知能力”が反映されている可能性がある。しかし同時に、メタ認知能力の獲得を想定しなくとも、教育的介入により単に「メタ認知は重要だ」と知ったことが、メタ認知に関する質問項目を社会的に望ましいと考えるようにさせ、測定値が向上したという説明も成り立つ。

末尾になるが、筆頭著者および第2著者は、直接観察することが困難であるメタ認知を、なんとかその目に収めようと試みてきた先行研究の学術的意義と貢献度について強調したい。本研究で取り上げたメタ認知尺度を開発した研究は、それぞれ当時の学術研究の最先端を開拓した偉大な巨人である。メタ

認知尺度の項目表現は、理科学習でメタ認知ができていない子どもの姿を研究者や理科教師に分かりやすく啓蒙し、さらなる研究・実践を進めるための巨人の肩であり続けた。本研究の結果は、これらのメタ認知尺度を開発した研究、またメタ認知の測定方法を追究してきた心理学研究と近年急速に発展した心理統計法の3体の偉大な巨人の肩の上に立つことで見えた地平である。本研究もまた、メタ認知と向き合う後続の理科教育学研究の足場として貢献できることを切に願う。

註

- 1) 以下では“理科学習”の達成指標として学業成績を使用することの問題点について考察する。筆記試験による学業成績では“真の理科の学力”の一部しか捉えられず、例えば、①生徒の動機づけなどの情意的側面、②教師の指導上の特徴などの要因が反映されないという問題が考えられる。

しかしながら、本研究では以下の3点を考慮し、学業成績を使用した。1点目は上記①の問題に関することである。例えば、子どもの動機づけは $d=0.47$ 、集中力・我慢強さ・積極性は $d=0.48$ 、教科に対する態度は $d=0.36$ の効果量で学業成績に影響することがメタ分析によって明らかになっている（Hattie, 2009a 西村訳, 2018）。そのため、生徒の情意的側面の状態はある程度学業成績に反映されていると考えられる。

2点目は上記②の問題に関することである。例えば、教師と学習者の関係は $d=0.72$ 、教師が学習者に質問することは $d=0.46$ 、探究的指導は $d=0.31$ 、問題解決的指導は $d=0.61$ 、学校全体で取り組む授業改善は $d=0.22$ の効果量で学業成績に影響することがメタ分析によって明らかになっている（Hattie, 2009b, c, d 大下・（一財）応用教育研究所・篠ヶ谷訳, 2018）。そのため、教師の指導上の特徴はある程度学業成績に反映されていると考えられる。また、6.3節で述べたように、現時点では教師要因によるメタ認知尺度と学業成績との相関の調整効果を明らかにした知見はなく、実際、研究3において確かな相関の調整は見出されていない。

3点目は先行研究の方法論を踏襲することの利点である。Ohtani & Hisasaka (2018) は学業成績として認知的領域の達成指標を用いた数多くの研究をメタ分析し、メタ認知と学業成績との正の相関を確認している。すなわち、メタ認知尺度が妥当にメタ認知を正しく捉えているのであれば、認知的領域のみを扱った学業成績でも確かな正の相関を示すのである。本研究ではOhtani & Hisasaka (2018) と同様の指標を用いることで、こうした強力なエビデンスと比較することを重視した。

上述した事柄を考慮し、本研究では、理科学習を取り巻く諸要因はある程度学業成績に反映されていると

捉え、全国学力・学習状況調査による学業成績には“真の理科の学力”が十分に反映されているとみなし、メタ認知尺度と理科の学業成績の相関係数を検討対象とした。

- 2) 全国学力・学習状況調査が“真の理科の学力”を“ある程度”反映するならば、その反映する割合に応じた相関の希薄化が生じる。

例えば、“真の理科の学力”が測定できたと仮定し、全国学力・学習状況調査は“真の理科の学力”と.70の相関(信頼性:.49)であるとする。メタ認知尺度の信頼性は研究1と研究2のの係数より.90とする。また、メタ認知尺度が“真の理科の学力”と.50の相関であるとする(目安として、on-line メソッドの相関: Ohtani & Hisasaka, 2018)。希薄化された相関係数は、真値との相関係数と信頼性の幾何平均との積であるので(狩野, 2002)。

$$0.50 \times \sqrt{0.49 \times 0.90} = 0.332 \dots$$

と求められる。すなわち、全国学力・学習状況調査が“真の理科の学力”をおよそ49%程度反映し、メタ認知尺度が“真の理科の学力”とon-line メソッドと同程度の相関があったとすると、全国学力・学習状況調査の得点と $r=.33$ 程度が検出される。同様の方法で計算すると、全国学力・学習状況調査と“真の理科の学力”の相関が.60(信頼性:.36)ならば $r=.28$ 程度、.50(信頼性:.25)ならば $r=.24$ 程度、.40(信頼性:.16)ならば $r=.19$ 程度の相関が検出されるはずである。

しかし、本研究の結果によるとメタ認知尺度が全国学力・学習状況調査の得点と.20以上の相関を持つ確率は小さかった。そのため、メタ認知尺度の妥当性が問題なのではなく、全国学力・学習状況調査が“真の理科の学力”を反映していないために本研究の結果が得られたという説は、全国学力・学習状況調査が“真の理科の学力”と約.40未満の相関しかなく、“真の理科の学力”を16%も反映していないという状況において説得力を持つ。そうであるならば、早急に全国学力・学習状況調査を改善する必要がある。

- 3) メタ認知尺度の妥当性は十分であり、全国学力・学習状況調査で測定される学力の側面以外とは十分な相関があるという説は、言い換えると、全国学力・学習状況調査で測定される理科の学力の側面はメタ認知と関連しないことを意味する。科学的思考にはメタ認知が重要であることを鑑みると、理科教育学として早急に全国学力・学習状況調査の妥当性を議論する必要がある。

附記

本研究は平成30年度日本理科教育学会四国支部大会、および第69回日本理科教育学会全国大会(静岡大会)で発表したデータに新たなデータと分析を加え、大幅に加筆したものである。

謝辞

調査協力校の生徒の皆様及び先生の皆様には、多大なご協力をいただきましたことを深く感謝申し上げます。また本研究の遂行にあたり中山留美子先生(奈良教育大学)よりご指導を賜りました。ここに感謝を申し上げます。

引用文献

- American Psychiatric Pub. American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed., DSM-5). Washington, DC: American Psychiatric Publishing. (アメリカ精神医学会 高橋三郎・大野裕(監訳)(2014)『DSM-5 精神疾患の診断・統計マニュアル』, 医学書院.)
- Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning?. *Metacognition and Learning*, 3(1), 39–58.
- Baumeister, R.F., Campbell, J.D., Krueger, J.L., & Vohs, K.D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological science in the public interest*, 4(1), 1–44.
- Crocker, J., & Wolfe, C.T. (2001). Contingencies of self-worth. *Psychological Review*, 108, 593–623.
- Dent, A.L., & Koenka, A.C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28(3), 425–474.
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and learning*, 3(3), 231–264.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175–191.
- Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, 34(10), 906–911.
- Gabriel, M.T., Critelli, J.W., & Ee, J.S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *Journal of Personality*, 62(1), 143–155.
- Hattie, J. (2009a). 4 The contributions from the student. In J. Hattie (Ed.), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. 39–60. London: Routledge. (ジョン・ハッティ 西村多久磨(訳)(2018)「第4章 学習者要因の影響」ジョン・ハッティ 山森光陽(監訳)(2018)『教育の効果: メタ分析による学力に影響を与える要因の効果の可視化』, 図書文化,

- 82-97.)
- Hattie, J. (2009b). 7 The contributions from the teacher. In J. Hattie (Ed.), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. 108-128. London: Routledge. (ジョン・ハッティ 大下卓司 (訳) (2018) 「第7章 教師要因の影響」ジョン・ハッティ 山森光陽 (監訳) (2018) 『教育の効果：メタ分析による学力に影響を与える要因の効果の可視化』, 図書文化, 137-153.)
- Hattie, J. (2009c). 9 The contributions from teaching approaches—part I. In J. Hattie (Ed.), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. 161-199. London: Routledge. (ジョン・ハッティ (一財) 応用教育研究所 (訳) (2018) 「第8章 指導方法要因の影響 I」ジョン・ハッティ 山森光陽 (監訳) (2018) 『教育の効果：メタ分析による学力に影響を与える要因の効果の可視化』, 図書文化, 154-204.)
- Hattie, J. (2009d). 10 The contributions from teaching approaches—part II. In J. Hattie (Ed.), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. 200-236. London: Routledge. (ジョン・ハッティ 篠ヶ谷圭太 (訳) (2018) 「第9章 指導方法要因の影響 II」ジョン・ハッティ 山森光陽監訳 (2018) 『教育の効果：メタ分析による学力に影響を与える要因の効果の可視化』, 図書文化, 205-251.)
- 久坂哲也 (2016) 「我が国の理科教育におけるメタ認知の研究動向」『理科教育学研究』, 第56巻, 第4号, 397-408.
- 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』, 第29巻, 第2号, 138-159.
- 加藤弘通・太田正義・松下真実子・三井由里 (2018) 「思春期になぜ自尊感情が下がるのか？—批判的思考態度との関係から—」『青年心理学研究』, 第30巻, 第1号, 25-40.
- 木下博義 (2010) 「ワークシート活用による子どものメタ認知促進に関する事例的研究：小学校第5学年「もののとけ方」を例に」『理科教育学研究』, 第51巻, 第2号, 11-19.
- 木下博義・松浦拓也・角屋重樹 (2005) 「観察・実験活動における生徒のメタ認知の実態に関する研究—質問紙による調査を通して」『理科教育学研究』, 第46巻, 第1号, 25-33.
- 国立教育政策研究所 (2018a) 「平成30年度全国学力・学習状況調査の調査問題・正答例・解説資料について」 Retrieved from https://www.nier.go.jp/18chousa/pdf/18mondai_chuu_rika.pdf (accessed 2019.06.08)
- 国立教育政策研究所 (2018b) 「平成30年度 全国学力・学習状況調査 報告書」 Retrieved from <https://www.nier.go.jp/18chousakekkahoukoku/report/data/18msci.pdf> (accessed 2019.06.08)
- 草場実・足達慶暢・鈴木達也 (2017) 「理科学習場面における高校生のメタ認知の実態に関する調査研究」『学習開発学研究』, 第10号, 75-81.
- 草場実・角屋重樹・森敏昭 (2012) 「メタ認知を活性化する観察・実験活動が高校生の実験観の変容に及ぼす効果：高等学校化学「化学反応と量の関係」を事例として」『日本教科教育学会誌』, 第34巻, 第4号, 29-38.
- 草場実・湯澤正通・角屋重樹・森敏昭 (2010) 「メタ認知を活性化する観察・実験活動が科学的知識の定着に及ぼす効果：高等学校化学「中和滴定」を事例として」『日本教科教育学会誌』, 第33巻, 第3号, 31-40.
- 松浦拓也・角屋重樹・岡田大爾・檜山芳之 (2002) 「観察・実験活動とメタ認知的技能の関係 (2)：電磁石作成課題の事例を通して」『科学教育研究』, 第26巻, 第5号, 350-357.
- 松浦拓也・木下博義 (2008) 「項目反応理論を用いたメタ認知尺度の構成に関する基礎的研究—理科学習における観察・実験活動を中心にして—」『日本教科教育学会誌』, 第30巻, 第4号, 1-7.
- 宮本樹・木下博義・網本貴一 (2015) 「高等学校化学における実験中のメタ認知育成に関する研究—酢酸エチルの合成実験を用いた実践を例として—」『理科教育学研究』, 第56巻, 第2号, 213-223.
- 文部科学省 (2016) 「幼稚園, 小学校, 中学校, 高等学校及び特別支援学校の学習指導要領等の改善及び必要な方策等について (答申) (中教審第197号)」 Retrieved from http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/_icsFiles/fieldfile/2017/01/10/1380902_0.pdf (accessed 2019.06.08)
- 中村晃 (2018) 「試験に対する自己評価と実際の点数の関連：自己愛と自尊感情を考慮に入れて」『千葉商大紀要』, 第56巻, 第1号, 89-98.
- 中山留美子・中谷素之 (2006) 「青年期における自己愛の構造と発達の变化の検討」『教育心理学研究』, 第54巻, 第2号, 188-198.
- Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, 13(2), 179-212.
- 大谷和夫・中谷素之 (2011) 「学業における自己価値の随伴性が内発的動機づけ低下に及ぼす影響プロセス—状態的自尊感情と失敗場面の感情を媒介として」『パーソナリティ研究』, 第19巻, 第3号, 206-216.
- 大谷和夫・中谷素之 (2013) 「中学生の学業領域における自己価値の随伴性が動機づけおよび学業達成に及ぼす影響プロセス—中間テストでの達成度を考慮に入れた縦断的検討」『パーソナリティ研究』, 第21巻, 第3号, 254-266.
- 岡田涼 (2009) 「青年期における自己愛傾向と心理的健康：

- メタ分析による知見の統合」『発達心理学研究』, 第20巻, 第4号, 428-436.
- 小塩真司 (2012) 「4 自分を好きになる」速水敏彦監修『コンピテンス 個人の発達とよりより社会形成のために』, ナカニシヤ出版, 30-37.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- 三宮真智子 (2008) 「第1章 メタ認知研究の背景と意義」三宮真智子 (編著) 『メタ認知 学習力を支える高次認知機能』, 北大路書房, 1-16.
- 三宮真智子 (2010) 「メタ認知は自己調整学習とどのように関係するのか」三宮真智子 (編著) 『教育心理学』, 学文社, 48-49.
- Schellings, G.L., van Hout-Wolters, B.H., Veenman, M.V., & Meijer, J. (2013). Assessing metacognitive activities: the in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European journal of psychology of education*, 28(3), 963-990.
- 清野樹恵・及川宏輝・古本温久・久坂哲也 (2018) 「ふきだし法を用いた理科の問題解答におけるメタ認知のオンライン評価」『日本科学教育学会研究会研究報告』, 第33巻, 第1号, 55-60.
- 鈴木誠 (1997) 「理科教育における学習意欲の構造に関する研究 (4) 一児童や生徒の自己効力感, 認知的方略のメタ認知, 及び社会的関係性の発達の变化について一」『日本理科教育学会研究紀要』, 第38巻, 第1号, 11-21.
- 鈴木誠 (1999) 「理科の学習場面における自己効力感, 学習方略, 学業成績に関する基礎的考察」『理科教育学研究』, 第40巻, 第1号, 11-23.
- 谷伊織 (2008) 「バランス型社会的望ましき反応尺度日本語版 (BIDR-J) の作成と信頼性・妥当性の検討」『パーソナリティ研究』, 第17巻, 第1号, 18-28.
- van der Stel, M., & Veenman, M.V. (2014). Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective. *European journal of psychology of education*, 29(1), 117-137.
- Veenman, M.V.J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt, & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis*. 75-97. Berlin: Waxmann.
- Veenman, M.V.J. (2012). Metacognition in Science Education: Definitions, Constituents, and Their Intricate Relation with Cognition. Zohar, A. and Dori, Y.J. (Eds.), *Metacognition in Science Education*. 21-36. Dordrecht, Heidelberg, London, New York: Springer.
- Veenman, M.V.J., Bavelaar, L., De Wolf, L., & Van Haaren, M.G.P. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, 29, 123-130.
- Veenman, M.V., Van Hout-Wolters, B.H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning*, 1(1), 3-14.
- 湯澤正通 (2008) 「科学的思考と科学理論の形成におけるメタ認知」三宮真智子 (編著) 『メタ認知 学習力を支える高次認知機能』 北大路書房, 131-149.

(2019年6月10日受付, 2019年10月9日受理)

Reconsidering Metacognition Measurement Questionnaires for Science Education: Limitations of Off-line Methods and Recommendations

Yuki HARADA^{1,2}, *Tetsuya HISASAKA*³, *Minoru KUSABA*², *Makoto SUZUKI*⁴

¹ Research Fellow of the Japan Society for the Promotion of Science

² Faculty of Education, Kochi University

³ Faculty of Education, Iwate University

⁴ Graduate School of Science, Hokkaido University

SUMMARY

Science education researchers and teachers consider cultivating students' metacognitive abilities to be important. Therefore, measurement questionnaires have been developed to take stock of these abilities. However, recent studies have pointed out that it is challenging to measure metacognition via self-reported measurement methods (off-line method) such as the questionnaire method. This study aimed to confirm whether questionnaires can truly measure metacognition in science learning. We verified the convergent validity and discriminant validity of metacognition measurement questionnaires developed for science education thus far. The results of our literature survey and analysis revealed that no substantial positive correlation between metacognition measurement questionnaires and academic achievement in science had yet been found. Genuinely positive correlations, however, were found between metacognition measurement questionnaires and narcissism or social desirability bias. Neither convergent validity nor discriminant validity was supported. Based on the above results, in this paper we further consider the limitations of metacognition measurement methods in current science educational research and corresponding strategies in the future.

<Key words> metacognition, measurement, validity, questionnaire, bayesian estimation

附録 本研究に使用したメタ認知尺度と弁別的妥当性の検討に使用した心理尺度

項目

メタ認知尺度

鈴木 (1997)	<p>わたしは理科の授業がわからないとき、そのわけがわかります 理科の成績が悪いとき、わたしは何がダメなのかわかります わたしは理科の授業の内容が、やさしいかむずかしいかわかります わたしは、理科で何を学習しているかわかります</p>
松浦・木下 (2008)	<p>わたしはこれから何を調べるのか、考えるようにしています 今までに習ったことを思い出しながら、予想を立てるようにしています 予想を確かめることができる実験方法になっているかどうか、確認するようにしています クラスのみんなの話し合いで友達の意見を聞いて、自分の意見を考え直すことがあります 次に何をやるのか考えながら、観察や実験をするようにしています 大事なところはどこか、考えるようにしています 何を確かめるのか考えながら、観察や実験をするようにしています 観察や実験がうまくいっていると感ずることがあります わからなくなったら、少し前にもどって、もう一度考えるようにしています 観察や実験の結果が予想どおりにならなかつたら、やり直すようにしています 自分は何を調べたのか、ふり返るようにしています グループの話し合いで、友だちの意見と自分の意見を比べながら聞くようにしています グループで話し合いをしていると、自分の意見がまとまることがあります クラスみんなの話し合いで、友だちの意見と自分の意見を比べながら聞くようにしています クラスみんなでお話し合いをしていると、自分の考えがまとまることがあります 先生の説明を聞いていると、自分の考えがまとまることがあります</p>
草場ら (2017)	<p>私は、理科の勉強で、新しく学んだことを、自分なりに整理する方法を知っています 私は、理科の授業が終わると、何が理解できて、何が理解できなかったのか、振り返るようにしています 私は、理科の勉強では、自分に合った学習方法をいろいろと試すようにしています 私は、理科の勉強は、どのような段取りで行えば、効率的に進むのかを知っています 私は、理科の問題を解くとき、どのような知識が必要なのかを考えるようにしています 私は、理科の授業で、分からないところがあれば、一人では考えず、先生や友だちに聞くようにしています 私は、理科の勉強は、目標を持って取り組むことで、より多くの知識が身に付くことを知っています 私は、理科の問題を解いている最中に、正しく解けているのか、こまめに確認するようにしています 私は、理科の問題を解く前に、問題文をよく読み、理解してから解くようにしています 私は、理科の問題を解くとき、自分なりにうまく解く方法を知っています 私は、理科の問題が解けなかった時、なぜ解けなかったのか、その理由を考えるようにしています 私は、理科のテストがあるとき、事前に計画を立ててから勉強を進めるようにしています</p>
自尊心	<p>自分には、よいところがたくさんあると感じています 私は、いつも失敗ばかりしていると思っています わたしは、自分自身にだいたい満足しています ときどき、自分は役に立たない人間だと思ふことがあります</p>
自己愛傾向 (誇大性)	<p>わたしには、持って生まれたすばらしい才能があります わたしは、周りの人からもっと高く評価されてもよい人間だと思います 自分自身では、要領もよいし、うまく判断のできるような賢さも備えていると思います わたしは、他に並ぶ人がいないくらい、特別な存在です</p>
社会的望ましき反応バイアス (自己欺瞞)	<p>わたしは自分で決めたことを後悔しません 自分で決心したことが他人の意見で変わることは滅多にありません わたしのことを、他人が本当にどう思っているのか気になりません わたしは、自分の人生を完全に思い通りに進めています わたしが感じた、他人の第一印象はよく当たっています</p>