

BEVI と IDI の比較

— その基本的特徴と妥当性に関するエビデンス —

永井 敦

The present paper compares the *Beliefs, Events, and Values Inventory* (BEVI) and the *Intercultural Development Inventory* (IDI) to explore similarities and differences in terms of basic characteristics and evidence related to their validity as a psychological measure. It was found that the two measures are largely similar in their basic characteristics but differ in the number of items included, how long it takes to complete them, and the nature of sample on which these measures were developed. There was a relatively good amount of validity evidence for both measures. The paper concluded that study abroad program administrators should choose, for program evaluation purposes, an appropriate measure that best fits the purpose or expected learning outcomes of their study abroad program.

1. 研究の背景と目的

グローバル人材育成という政策課題のもと、日本人による海外留学には相当の資源が投資されてきた。海外に留学する学生数は着実に増加し、平成 29 年度には 66,058 人にのぼった（日本学生支援機構、2019 年）。

しかしながら、我が国では河合塾（2018）が指摘するように「留学を通して学生がグローバル人材としてどのように成長しているのかといった質の観点からの効果検証は、必ずしも十分に手掛けられてこなかった」（p.1）。

このような中で、近年、日本の大学において留学の効果測定に関する関心が急速に高まっている。例えば日本私立大学連盟が刊行する隔月誌『大学時報』の 2018 年 5 月号で「海外留学体験の効果測定に対する取り組み —海外短期派遣プログラムを中心に—」という特集が生まれ、いくつかの代表事例が取り上げられている。

これらの取り組みは歓迎すべきものであるが、永井（2018）が指摘しているように、日本の高等教育機関が実施する留学の効果測定ツールは、留学プログラム運営担当者が独自に開発したものが多く、教育測定学の知見（例えば Brennan, 2006）や標準的な尺度構成の手続き（例えば Devellis, 2016）が踏まえられておらず、測定ツールの信頼性や妥当性が十分に検討されているとは言えない。そのような状況の中で、上で紹介した『大学時報』の特集で広島大学が紹介する、米国 James Madison 大学の臨床心理学者である Craig N. Shealy 博士が中心となって開発した *Beliefs, Events, and Values Inventory*（略称 BEVI）はそれらの尺度構成の手続きをふまえ、信頼性と妥当性の高い心理尺度として、留学の効果測定への応用が各方面から期待されており、また、

実際に国内の複数の大学でも実際に導入されている¹。また、平成 28 年 12 月 8 日には東京にて、文部科学省後援による「留学の学習成果分析 (BEVI-j) シンポジウム」が開催されている²。また、「平成 30 年度大学の世界展開力強化事業～COIL 型教育を活用した米国等との大学間交流形成支援～」においては、BEVI がその教育成果を測定するツールとして導入されていることからその注目度の高さがわかるであろう。

しかしながら、この新たな流れの中において、筆者は実際に BEVI を高等教育機関で日常的に利用する者として、今後国内の留学プログラム担当者が BEVI を無批判的に、留学の効果測定のための万能ツールのよう捉えてしまうことを懸念している。留学の効果測定目的で応用可能な心理尺度 (本稿ではテストと読み替えて良い) は、世界的に見てもかなりの数があり (例えば異文化間能力について Fantini (2006) は 50 以上の測定ツールを紹介している)、BEVI はあくまでその中の 1 つにしかすぎない。だが、それら心理尺度の性質を正しく理解して、各留学プログラムに最適なツールを比較考量するためには (測定「内容」の背後にある理論の理解は前提として) 統計学に基づく心理測定・教育測定の専門的な知識を必要とする。そのため、それらの訓練を必ずしも受けていない留学プログラムの担当者には、端的に言って、荷が重い作業であろう。

本稿の目的は、上記の問題意識のもと、大学等の留学プログラム担当者の、BEVI の心理尺度としての性質の理解を深め、その応用的価値の相対化を図ることである。その目的達成のため、BEVI と同様に留学の効果測定分野でよく知られている *Intercultural Development Inventory* (略称 IDI) を取り上げ、BEVI と比較、両者の特徴を検討する。

以下では、まず両尺度の基本的な特徴を、両尺度の専用 HP で公開されている情報及び関連文献に基づいて整理し、比較する。その後、(心理) 尺度にとって最も重要な要素である妥当性に関するエビデンスの視点から考察する。

2. BEVI 及び IDI の心理尺度としての基本的特徴

表 1 は BEVI 及び IDI の基本的な特徴をまとめたものである。

どちらの尺度も、関連する主要な文献は 2000 年代以降に出版されている。背景理論そのものについての説明は本稿の趣旨ではないため割愛するが、それぞれ文献の特定が可能な、明示的な背景理論 (EI 理論と DMIS 理論) を基盤に持っていることは重要な点である。測定内容については、BEVI が「人生の経験、文化及び文脈との関係における我々の価値、信念及び世界観」を測定するようにデザインされているのに対し、

¹ 本稿執筆時点では、BEVI データを用いて書かれた唯一の研究論文として永井 (2018) がある。同論文では BEVI の簡単な紹介も行われているため、興味のある読者は参考にされたい。

² <https://www.hiroshima-u.ac.jp/news/36769> を参照。

IDI は「文化的視点を変化させ、文化的差異や共通点に適切に行動を合わせる能力」を測定するための尺度である。言い換えれば BEVI は価値や信念といったより幅広い概念を測定するが、IDI は異文化理解に関わる能力の測定に特化していると言える。いずれの尺度も質問紙型(オンラインで回答可能)のフォーマットで作成されており、回答者から 4 または 5 段階の反応(例えば刺激文に対して“全く同意しない—同意しない—同意する—完全に同意しない”)のような選択肢が与えられる)を引き出すデザインである。各尺度に含まれる項目数は、BEVI が 185 項目に対し、IDI は 50 項目であり、この差は回答に要する時間の差にもつながっている。本稿執筆時点では BEVI は英語と日本語の 2 言語で利用可能だが、IDI は日本語を含む 17 言語に対応している。尺度の実施者には、アドミニストレーターとして特別な研修(有料)が必要である。尺度の使用料は BEVI が年間ライセンス制であるのに対し、IDI は回答者一人当たりで料金が発生する。尺度開発時に用いられたサンプルについては BEVI が主に大学生を対象としたのに対し、IDI は年齢、文化、言語的に非常に多様な人々のデータに基づいて開発されている。尺度の信頼性(本稿においては Cronbach の信頼性係数で測定される内的一貫性を意味する)については、下位尺度によってばらつきが見られる。 α 係数は一般に「信頼性の下限値なので、0.8 以上の値が得られた場合は問題ない」(村上, 2006, p.37) とされる。実際にはそれぞれの尺度の α 係数の最低値が .6 程度になっており、特に IDI は信頼性係数の最高値も 0.8 を満たしていないが、IDI 開発者の一人である Hammer (2011, p.482) は“文化的、言語的に多様なサンプルであることを考慮すると、これらの信頼性は十分なレベルの項目間信頼性を持っていると言える”としている。

表 1. BEVI 及び IDI の基本的特徴

尺度	BEVI (Short Version ³)	IDI
典拠	Wandschneider et al. (2015) ; Shealy (2016) ; BEVI (2019)	Hammer et al. (2003) ; Hammer (2011) ; IDI (2019)
開発年 ⁴	2013	2003
背景理論	EI 理論 (Shealy, 2004)	DMIS 理論 (Bennet, 1986)
フォーマット	質問紙型 (評定尺度法: 4 件法)	質問紙型 (評定尺度法: 5 件法)

³ 現在国内で用いられている BEVI はその短縮版であるため、本稿で BEVI に言及する際は、それは BEVI 短縮版を意味する。

⁴ 両尺度に関する文献が開発(完了)年について明示的に言及していなかったため、これは筆者が公開されている文献をもとに推定した数字である。

測定対象（テーマ）	人生の経験、文化及び文脈との関係 での我々の価値、信念及び世界観	文化的視点を変化させ、文化的差異 や共通点に適切に行動を合わせる 能力
実施形態	オンライン	オンライン/オフライン
項目数	185（+40の属性項目） *3つの質的項目あり	50（+10の属性項目） *4つの質的項目あり
回答に要する時間	25 ～ 30分	15 ～ 20分
利用可能言語	2言語（日本語含む）	17言語（日本語含む）
テスト実施者の研修	必要（有料）	必要（有料）
使用料	年間ライセンス制 1~200人：1,000 USD 201~1000人：3,000 USD 1001~2000人：5,000 USD 2001~5000人：7,000 USD 人数無制限：10,000 USD	一人当たり 11 USD *教育機関での、所属学生を対象と した場合
尺度開発のサンプル	主に大学生（サンプル全体の 96.7%）	高校生から壮年あたりまでの多様 な人々
信頼性（ α 係数）	.61 ～ .90	.66 ～ .79

3. BEVI 及び IDI の妥当性に関するエビデンス

上では BEVI と IDI という 2 つの尺度の基本的特徴を概観してきた。ここでは「心理測定上最も重要な概念」（村上, 2006, p.52）である「妥当性」に焦点を当てて両尺度を考察していく。

教育学や心理学での尺度は、ある特定の構成概念を測定するために開発されるが、尺度が「測定していると主張している内容を本当に測定できているかどうか」（Kelly, 1927, p.14）という妥当性の問題は、実証的に検証されなければならない。心理測定の分野で頻繁に引用される Messick（1995, p.741）によれば「妥当性とはテスト得点もしくは他のアセスメントにもとづく解釈と行動の適切さと十分さについて、それらを支える実証的証拠や理論的根拠がどの程度あるかに関する、総合的かつ評価的な判断」⁵である。この定義から、妥当性は（有るか無いかではなく）程度の問題であること、妥当性は（単一のエビデンスで構成されるのではなく）多様なエビデンスに基づいて

⁵ “Validity is an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment.”

総合的に評価されることがわかる。Messick (1995) による妥当性の概念 (Messick, 1989 も参照されたい) はその後いくつかの重要な研究的発展を経て (例えば Brennan, 2006)、*Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) に反映され、現場で実際に利用されている。本稿はこの *Standards* で定義されている 5 つの妥当性エビデンスの観点から BEVI と IDI を考察、比較する。

1 つ目の妥当性エビデンスは尺度の「内容」 (evidence based on test content) に関するものである。これは尺度を構成する項目が、狙いどおりの測定対象を測定しているかどうかに関わるエビデンスである。このエビデンスは典型的には、測定内容 (背景理論) に関する専門家を含むグループが項目内容を精査することで提供される。BEVI も IDI もどちらも専門家によるレビューが行われているが、BEVI については特に実際のカウンセリングのクライアントによる発話データを基に測定項目を開発している点が特徴的である。

2 つ目の妥当性エビデンスは構成概念と、(尺度) 回答者のパフォーマンスや実際に従事する「反応プロセス」 (evidence based on response processes) との一致度に関わるものである。一般に回答者の中で実際にどのような心理過程が生起しているかは観察できないため、思考発話法 (think aloud) などの手続きを用いて、証拠を集めていく。BEVI も IDI も、この点での明示的に示されているエビデンスは無く、今後の研究課題と言える。

3 つ目の妥当性エビデンスは内部構造 (evidence based on internal structure) に関わるもので、因子分析や構造方程式モデリングなどの手法を用いて、尺度が背景理論の予測するような (因子) 構造を示すかどうかを統計的に検証することで妥当性を示す。BEVI も IDI も統計的手法を用いて内部構造に関わるエビデンスを提出している。

4 つ目は測定対象の構成概念と他の (例えば類似する、または異なる) 変数 (構成概念) との関係 (evidence based on relations with other variables) に関わるエビデンスである。例えば、背景理論上類似すると考えられる構成概念同士は、相関係数が高いはずであり (収束的証拠)、逆に異なると考えられる構成概念間では、相関係数が低くなるはずである (弁別的証拠)。BEVI も IDI も他変数との関係について限定的ながら、検証を行っている。

5 つ目の妥当性エビデンスは、尺度の得点の解釈が尺度の実施目的 (例えばクラス分けや診断) に照らし合わせて適切なものであるかに関わるものである (evidence based on the consequences of testing)。BEVI も IDI もいずれも得点の適切な解釈及び利用に関する論文を公表しており、また、定期的にワークショップを開催することでそれぞれの尺度の誤用を防ぐ努力をしている。尺度の得点の過大解釈の問題に関わることとして、BEVI は英語圏の文脈で、大学生を主なサンプルとして開発された経緯

から、日本語版 BEVI（英語版の内容を忠実に日本語に翻訳したもの）の項目にやや英語圏の文化に偏った内容が散見される。一方 IDI は、尺度構成段階で非常に多様な背景（年齢や文化）を持つサンプルに基づいて開発されたため、その点では比較的文化的バイアスが低い。

表 2. BEVI 及び IDI の妥当性に関するエビデンス

	BEVI	IDI
内容	実際のカウンセリングのクライアントによる信念・価値に関する発言を臨床研究者、実践家（カウンセラー等）が内容を精査し、項目を作成	文化差に関する経験を持つ多様な人々にインタビュー調査を行い、そのトランスクリプトを研究者チームが DMIS に基づいて評価し、評価の一致度に基づきつつ、項目を作成
反応プロセス	なし	なし
内部構造	探索的・確認的因子分析及び項目反応理論によって、理論から予測される因子構造との整合性を確認	探索的・確認的因子分析によって、理論から予測される因子構造との整合性を確認
他構成概念との関係	<i>Environmental concerns</i> (Patel, 2008) 及び <i>intercultural awareness</i> (Reisweber, 2008) などの、グローバル意識との関連概念を予測	<i>World-mindedness Scale</i> (e.g., Wiseman, Hammer, & Nishida, 1989) 及び <i>Intercultural Anxiety Scale</i> (Stephan & Stephan, 1985) と理論から予測されるとおりの相関傾向を確認
テストの結果	<ul style="list-style-type: none"> ・現場でのテスト結果の適切な使用・応用について論文を公表 ・定期的なワークショップ・セミナーの実施 ・英語以外でのテストは文化背景のバイアスがややあり 	<ul style="list-style-type: none"> ・現場でのテスト結果の適切な使用についてのケーススタディを多数実施、報告 ・定期的なワークショップ・セミナーの実施 ・テストバイアスは低い

4 まとめ

BEVI と IDI とともに、標準の尺度構成の手続きをふまえて作成されており、妥当性に関するエビデンスも比較的充実していると言える。

尺度の基本的特徴についても類似点が多いが、実務的な視点からは尺度の項目数は少ない方が実施や回答が容易である。BEVI の 185 項目は IDI の 3 倍以上の長さであり、

回答者への負担がやや大きい。尺度開発時に依拠したサンプルの多様性の点でも BEVI と IDI は大きく異なる。しかし、これらの違いも結局は測定内容（目的）の差に帰着するため、最後にその点について言及し、本稿を締めくくりたい。

BEVI は信念や価値を含む幅広い心理的変化を測定するが、IDI は異文化発達の測定に特化している。したがって留学プログラムの効果測定のための最適なツールを選択するという文脈においては、プログラム担当者はまず担当の留学プログラムが参加者のどのような知識、態度、能力面での成長を期待してデザインされているかを考慮しなくてはならない。そして、その目的によりふさわしい測定ツールを選択する必要がある。例えば現在、日本政府や各高等教育機関の留学プログラムの多くが将来の「グローバル人材」として留学経験者に異文化間能力（Intercultural Competence）面での成長を期待していることをふまえると、IDI が最適な効果測定ツールとして選ばれるかもしれない⁶。もしその留学プログラムが価値観や信念といったより幅広い側面での成長を期待している場合は BEVI が（例え項目数が多くても）、より目的にふさわしい効果測定ツールとして選ばれることになるだろう。このように留学の効果測定ツールは、プログラムの目的に照らしてその使用の適切さを相対的に判断されるべきであり、先にツールありきの話ではないことを強調しておきたい。

本稿では留学プログラム担当者の、近年注目されている BEVI への理解を深め、また、その応用的価値の相対化を図ることを目的として、IDI という別の尺度とともに、その基本的特徴及び妥当性に関するエビデンスについて比較考察を行った。本小論が我が国での今後の留学効果測定の議論において、関係者の測定ツールへの意識向上の一助となれば幸いである。

参考文献

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (2014) . *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Beliefs, Events, and Values Inventory. (2018) .Retrieved February 28, 2019, from <http://thebevi.com/>
- Bennett, M. J. (1986) . A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations*, 10 (2) , 179–196.
- Brennan, R. L. (Ed.) (2006) . *Educational measurement* (4th ed.) . Westport, CT: American Council on Education and Praeger.
- DeVellis, R. F., (2016) . *Scale development: Theory and applications* (4th edition) . Thousand Oaks, CA: Sage Publications.

⁶ BEVI も異文化理解に関係する下位尺度として社会文化的開放性（Sociocultural Openness）が存在するが、下位尺度として全体の一部でしかなく、この測定目的においてはそれに特化している IDI の方がより信頼性と妥当性が高い。

- Fantini, A.E. (2006). *Assessment tools of intercultural communicative competence*. Retrieved February 28, 2019, from <http://federazioneil.org/documents/AppendixF.pdf>.
- Hammer, M. R. (2011). Additional cross-cultural validity testing of the Intercultural Development Inventory. *International Journal of Intercultural Relations*, 35, 474-487.
- Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). The Intercultural Development Inventory: A measure of intercultural sensitivity. *International Journal of Intercultural Relations*, 27, 421-443.
- 河合塾 (2018). 平成 29 年度 文部科学省 委託事業「日本人の海外留学の効果測定に関する調査研究」成果報告書.
- Kelley, T. L. (1927). *Interpretation of educational measurements*. New York: Macmillan.
- Intercultural Development Inventory. (2019). Retrieved February 28, 2019, from <https://idiinventory.com/>
- Messick, S. (1989). Validity. In: R. Linn (Ed.), *Educational measurement* (pp. 13-104). New York: American Council on Education & Macmillan Publishing.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- 村上宣寛 (2006). 『心理尺度のつくり方』 北大路書房.
- 永井敦 (2018) 「BEVI によるショート・ビジット型留学プログラムの効果分析：「グローバル人材」は育成できるのか？」『広島大学留学生センター紀要』第 22 号, pp.38-52.
- 日本私立大学連盟 (2018). 大学時報. 第 380 号 (5 月号). 一般社団法人日本私立大学連盟.
- Patel, R. (2008). Environmental beliefs, values, and worldviews: Etiology, maintenance, and transformation (Unpublished doctoral dissertation). James Madison University, Harrisonburg, VA, the United States.
- Reisweber, J. R. (2008). *Beliefs, values, and the development of intercultural awareness* (Unpublished doctoral dissertation). James Madison University, Harrisonburg, VA, the United States.
- Shealy, C. N. (2004). A model and method for “making” a C-I psychologist: Equilintegration (EI) theory and the Beliefs, Events, and Values Inventory (BEVI). [Special Series]. *Journal of Clinical Psychology*, 60 (10), 1065-1090.
- Shealy, C. N. (Ed.) (2016). *Making sense of beliefs and values*. New York: Springer Publishing.
- Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, 41, 157-175.
- 独立行政法人日本学生支援機構 (2019). 平成 29 年度協定等に基づく日本人学生留学状況調査.
- Wandschneider, E., Pysarchik, D. T., Sternberger, L. G., Ma, W., Acheson, K., Baltensperger, B., Good, R. T., Brubaker, B., Baldwin, T., Nishitani, H., Wang, F., Reisweber, J., & Hart, V. (2016). The forum BEVI project: Applications and implications for international, multicultural, and transformative learning. In C. N. Shealy (Ed.), *Making sense of beliefs and values: Theory, research, and practice* (pp. 407-484). New York, NY, US: Springer Publishing.
- Wiseman, R. L., Hammer, M. R., & Nishida, H. (1989). Predictors of intercultural communication competence. *International Journal of Intercultural Relations*, 13, 349-370.