

力なき者は一発勝負に賭けるべきか? (その2)

—シミュレーション実験からわかったこと—

Should lower-ability players bet on a one-shot match? (part 2) :

Findings from simulation experiments

竹網 誠一郎*、千葉 龍一**、迫井 紀子**、今井 真***、
松尾 侑佳**、渡辺 千聖**、濱田 晶子**

TAKETSUNA Seiichiro, CHIBA Ryuichi, SAKOI Noriko, IMAI Shin,
MATSUO Yuka, WATANABE Chisato, HAMADA Akiko

問題と目的

中学・高校の生徒は様々な学力試験を経験している。その際、生徒は事前に準備をして試験に臨んでいる。試験の結果、真の学力に近い得点を取る生徒がいる一方、準備していた箇所や得意な分野が多く出題されたことで真の学力以上の得点を取る生徒や逆に真の学力よりも低い得点を取る生徒もいる。心理統計学では、学力試験の測定値（成績）＝真の値＋誤差で表わされる。つまり、測定値には様々な要因によって生じる誤差が含まれており、必ずしも真の値ではないのである。竹網・土井・鎌原（2016）は試験得点と真の学力の関係について検討した。大学生47人に対し連合学習形式の暗記課題50ペアを印刷した課題用紙を配布し、1週間後に50題から10題を出題する小テストを実施すると告げた。1週間後、予告していない50題全問のテストを課した後、予告していた10題の小テストを実施した。この手続きにより、全問を出題した50題テストの得点は参加者の真の学力を測定していることになる。全体的に見れば、真の値（50題の得点）と測定値（10題からなる小テストの得点）の相関は高かった。しかし、参加者の2つの得点を個々の事例として丹念に調べてみると、真の学力が16～20点（50題中の3～4割）の参加者5名の測定値（小テスト得点）が2点～5点（2割～5割）というケースがあった。小テストにどの問題が出題されたかによって、測定値に誤差が生じていたのである。

竹網・稲葉・黒住（2018）は、課題遂行能力を測定しようとするのではなく、あらかじめ能力が異なる5名集団を実験的に作り出し、その後2とおりの対戦方法（1つは1回勝負、もう1つは10回勝負）で5名を競わせるという方法を用いた。実験には大学生60名（1集団5名×12集団）が参加し、以下の手順で実験が進められた。①各集団の成員5名は割り当てられた能力に応じて、1位者はトランプカード52枚の内の18枚を、2位者は14枚を、3位者は10枚を、4位者は6枚を、そして5位者は残り4枚すべてを獲得した。②実験者が実験者用の52枚のカードから1枚のカードを引き、それと同一のカードを持っている参加者がそのセッションの勝者となる旨を教示した。③最初に1回勝負を行った。実験者が52枚のカードから1枚のカードを引き、その当たりカードを持っている者が1位者となった。④その後、同じ能力差のまま10回勝負を行った。実験者は52枚のトランプカードから

* 学習院大学文学部心理学科教授、** 学習院大学大学院人文科学研究科臨床心理学専攻博士前期課程1年、*** 学習院大学大学院人文科学研究科心理学専攻博士後期課程1年

順次10枚のカードを引き、10回中の最多正解者が1位者となった（同点者が出た場合は該当者だけで決勝戦を行い、1位者1名を確定させた）。分析の結果、10回勝負では能力の高い1位群と2位群が1位者になることが多かった。しかし、1回勝負では、3位群と4位群が1位群・2位群以上に好成績だった。このことから、力なき者（課題遂行能力が低い者）は1回勝負、つまり一発勝負に賭けるべきだという結論が示された（先行研究の勝率データは本論文の図1および図2に示されている）。

しかしながら、竹綱ら（2018）が採用した実験手続には、生態学的妥当性上の問題があった。課題遂行の能力差を作り出すために、52枚のカードを1位者～5位者に所定の枚数を配分したため、実験者が示す当たりカードを持っている正解者はいつも1人しかいなかった。この状況は教室での競争場面を忠実に再現していることにはならない。何故なら、通常、教室では1つの問題の正解を複数の生徒が知っていることの方が自然だからである。より一般化した知見を得るためには、生態学的妥当性を高める必要がある。また、すべての試行で正解者が1人だけしかないという実験手続には、1回勝負における能力下位群のパフォーマンスを過大評価する危険性もあった。以上の問題点を解消するために、本研究では52枚のカードを5名に配分するのではなく、トランプカードを5組用意し、参加者は自分の順位に応じた枚数を52枚のカードから引く手続を採用することにした。さらに、先行研究では複数の同点者がいた場合、決勝戦を行って1位者を1人に絞る手続を用いていた。教室場面での試験において同点者がいることは当たり前のことである。本研究ではこの手続も修正し、1位該当者全員を1位者とした。本研究の実験1の目的は竹綱ら（2018）の実験手続きを修正し、追試することである。そして、10回勝負と1回勝負の2つの対戦方法が参加者の課題遂行能力をどの程度正確に把握しているかを検討することである。

竹綱ら（2018）の実験手続は知識量あるいは記憶量を問う試験であり、学校場面では漢字書き取りテストがその典型例と言える。しかし、定期試験や入学試験といった生徒にとって重要な学力テストでは、知識だけではなく技能、理解度や応用力を含めた総合的な学力を測定する試験を実施している。仮に、数学の総合的な学力が4つの側面（知識、技能、理解力、応用力）の合計点で捉えられるとすると、それぞれの側面を測定できる試験問題を作成するだろう。本研究の実験2の目的は、4つの側面からなる総合的な学力を測定する実験的課題を考案し、10回勝負と1回勝負の2つの試験方法が参加者の課題遂行能力をどの程度正確に把握しているかを検討することである。

本研究は、シミュレーション実験によってデータを収集した。集団成員（参加者）の能力差は実験の最初にランダムに決定された、その後、10回勝負および1回勝負を割り当てられた能力水準のまま実験課題に取り組んだ。次の集団も同様の手続で能力差がランダムに決定され、同様に実験課題に取り組んだ。12集団間のデータは互いに独立しているため、データ分析も独立したものとして扱った。

実験1

目的

実験1の目的は、竹綱ら（2018）の研究の生態学的妥当性を高めるように実験手続きを修正し、彼らの研究を追試することである。

方法

（参加者）5人1組による集団実験を、12集団を対象に実施した。

（能力差の決定方法）参加者にサイコロ6個を振ってもらい、その合計点の高い順に本実験の能力順位（1位～5位）を割り当てた。1位者は52枚から18枚のカードを獲得した。

2 位者～5 位者も52枚のカードから、2 位者は14枚、3 位者は10枚、4 位者は6 枚、そして5 位者は4 枚のカードを獲得した。この枚数がこの課題における彼らの学力の真の値を表わしている。

(課題) 実験者が実験者用52枚のカードから1枚のカードを引き、それを持っている参加者が正解者となる。10回勝負では、実験者は52枚のトランプカードから順次10枚のカードを引き、各回で当たりカードを持っている参加者全員が正解者となった。そして、10回中の最多正解者が1 位者となり、同点の最多正解者がいた場合は該当者全員を1 位者とした。引き続き、1 回勝負のセッションに入った。実験者が52枚のトランプカードから1枚のカードを引き、それを持っている参加者が1 位者となった。同一のカードを持っている参加者が複数人いた場合は該当者全員を1 位とした。

結果

10回勝負の1 位者と1 回勝負の1 位者の人数は、表1 のとおりである。10回勝負の1 位者は能力上位群で占められていたのに対して、1 回勝負では全群から1 位者が出ていた。先行研究と本研究の10回勝負における各順位群の結果を比較できるように、2 研究の勝率の比較図を図1 として示した。同様に、1 回勝負における2 研究の勝率の比較図を図2 として示した。2 つの図から、実験手続の生態学的妥当性を高めたことで、10回勝負においても1 回勝負においても、本研究の方が竹綱ら(2018) よりも能力を正確に把握できていることがわかった。

表1 10回勝負と1 回勝負の1 位者の人数

課題遂行能力の順位	10回勝負の1 位者人数	1 回勝負の1 位者人数
1 位 (12名)	8 名	10名
2 位 (12名)	5 名	4 名
3 位 (12名)	2 名	4 名
4 位 (12名)	0 名	3 名
5 位 (12名)	0 名	2 名
計	15名	23名

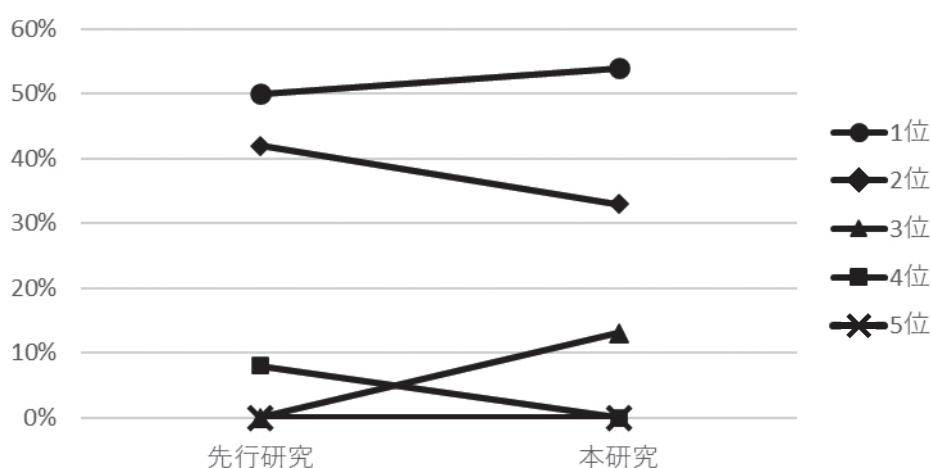


図1 10回勝負の場合の各順位の勝率

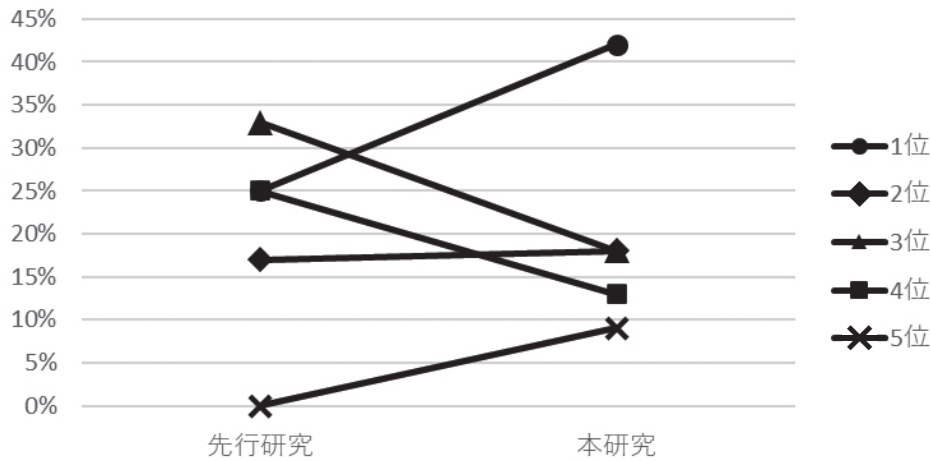


図2 1回勝負の場合の各順位の勝率

10回勝負において課題遂行能力1位群～5位群が1位になった人数と2位～5位になった人数を表2に示す。カイ2乗検定の結果、 $X^2(4) = 21.3, p < .01$ となった。残差分析の結果、遂行能力1位群は他の群に比べて1位者になりやすいこと、4位群と5位群は他の群に比べて共に1位者になりにくいことがわかった。一方、1回勝負において課題遂行能力1位群～5位群が1位になった人数と2位～5位になった人数を表3に示す。カイ2乗検定の結果、 $X^2(4) = 13.8, p < .01$ となった。残差分析の結果、能力1位群は他の群に比べて1位者になりやすいことが確認されたが、2位群から5位群は1位者になりやすくもなりにくくもないということがわかった。この結果は、力なき者は一発勝負に賭けるべきだという竹綱ら(2018)の結論を支持した。

生態学的妥当性を高めるために実験手続を修正したことで、10回勝負も1回勝負も参加者の能力差を先行研究よりも正確に把握できるようになった。特に、1回勝負の判定力はかなり改善したと判断できる。しかしながら、10回勝負による測定精度の高さは1回勝負のそれを圧倒しており、遂行能力を正確に測定するためには10回勝負の方が優れた選抜方法であることがわかった。

表2 10回勝負の1位者と2位～5位の人数

課題遂行能力の順位	10回勝負の1位者人数	10回勝負の2位～5位者人数
1位 (12名)	8名	4名
2位 (12名)	5名	7名
3位 (12名)	2名	10名
4位 (12名)	0名	12名
5位 (12名)	0名	12名
計	15名	45名

表3 1回勝負での1位者と2位～5位の人数

課題遂行能力の順位	1回勝負の1位者人数	1回勝負の2位～5位者人数
1位 (12名)	10名	2名
2位 (12名)	4名	8名
3位 (12名)	4名	8名
4位 (12名)	3名	9名
5位 (12名)	2名	10名
計	23名	37名

実験2

目的

本実験では、知識、技能、理解力と応用力といった諸側面の学力を総合的に捉える学力を想定し、実験手続を考案した。複数の能力を加算して算出した学力得点の分布は平均値の頻度数を頂点とし、それを中心に正規分布の形状になると考えられる。このような得点分布を示す学力を測定する道具として、本実験では6面サイコロを用いることにした。4つの諸側面それぞれを4個のサイコロで測定し、その合計点を総合的な学力と定義する実験方法を考案した。Barford (1985) は、真の値 X のユニークな特徴は (1) X はすべての測定値の平均値である、(2) X は最大頻度の値である、(3) X は中央値である、と述べている。4つのサイコロの合計点の平均値14点はこの3条件を満たしていることから、この値が真の学力を表していると判断した。

本実験では、参加者の課題遂行能力差を作り出すために、1位者群の参加者は4個のサイコロの目の合計に+2点を、2位者群には+1点を、3位者群は±0点を、4位者群は-1点を、5位者群は-2点を加算することにした。この結果、各群の参加者の学力の真の値は、上位から順に16点、15点、14点、13点と12点となる。真の値(平均値)は各群で異なるものも、標準偏差は5群とも同じ値となる。実験1では竹綱ら(2018)の方法に倣い12集団で実験を行ったため、カイ2乗検定を行う際に1つのセルの人数が少ない(期待値の値が小さい)という分析上の問題が生じていた。実験2ではこの問題を解消するために、24集団を対象に実験を行うことにした。

実験2の目的は、総合的な学力を測定できる実験課題において、10回勝負と1回勝負の2とおりのテスト方法が能力差をどの程度正確に測定しているかを検討することである。

方法

(参加者) 5人1組による集団実験を、24集団に対して行った。

(課題遂行能力の順位決定方法) 集団成員5名にサイコロ6個を振ってもらい、その合計点の高い順に本実験の課題遂行能力順位(1位~5位)を割り当てた。1位者は4つのサイコロを振って出た目の合計点に2点を加した得点を、2位者は出た目の合計点に1点を加した得点を、3位者は出た目の合計点を、4位者は出た目の合計点から1点を減した得点を、5位者は出た目の合計点から2点を減した得点をテスト得点とした。

(課題) 4つの側面を含む総合的な学力を測定するという考え方にも基づいて、実験参加者は4個のサイコロを振ることで課題遂行を行った。10回勝負では、各回の最高得点者が勝者となり、10回中の勝ち数最多者を1位者とした。勝ち数最多者が複数いた場合は該当者全員を1位者とした。引き続き1回勝負を行った。1回勝負では、最高得点者が1位者となった。最古得点者が複数いた場合は該当者全員を1位者とした。本研究も研究1と同様に、同一集団によるシミュレーション実験によってデータを収集した。

結果

10回勝負の1位者と1回勝負の1位者の人数は、表4のとおりである。10回勝負の1位者は能力上位の2群でほぼ占められており、5位群から1位になる者は1人もいなかった。しかし、1回勝負では4位群の1/6(17%)、5位群の1/8(13%)が1位者になっていることがわかった。

表4 10回勝負と1回勝負の1位者の人数

課題遂行能力の順位	10回勝負の1位者人数	1回勝負の1位者人数
1位 (24名)	13名	11名
2位 (24名)	12名	7名
3位 (24名)	3名	5名
4位 (24名)	2名	4名
5位 (24名)	0名	3名
計	30名	30名

表5 10回勝負の1位者と2位～5位者の人数

課題遂行能力の順位	10回勝負の1位者人数	10回勝負の2位～5位者人数
1位 (24名)	13名	11名
2位 (24名)	12名	12名
3位 (24名)	3名	21名
4位 (24名)	2名	22名
5位 (24名)	0名	24名
計	30名	90名

表6 1回勝負の1位者と2位～5位者の人数

課題遂行能力の順位	1回勝負の1位者人数	1回勝負の2位～5位者人数
1位 (24名)	11名	13名
2位 (24名)	7名	17名
3位 (24名)	5名	19名
4位 (24名)	4名	20名
5位 (24名)	3名	21名
計	30名	90名

10回勝負における課題遂行能力1位者～5位者が1位になった人数と2位～5位になった人数を表5に示す。カイ2乗検定の結果、 $X^2(4)=32.4, p<.01$ となった。残差分析の結果、能力1位者群は他の群に比べて1位になりやすいこと、能力4位者群と5位者群は他の群に比べて共に1位になりにくいことがわかった。1回勝負における課題遂行能力1位者～5位者が1位になった人数と2位～5位になった人数は表6のとおりである。カイ2乗検定の結果、 $X^2(4)=8.89, p>.05$ となり、5%水準では群間に有意差がみられなかった。しかし、事後検定の残差分析では、課題遂行能力1位群だけが5%水準で有意に他の群よりも1位になりやすいことが確認された。

総合的な学力を測定しようとした本実験においても、10回勝負の測定精度が極めて高いことが確認された。一方、1回勝負の測定精度は低く、カイ2乗検定ではどの群も1位者になりやすくもなりにくくもないという結果だった。しかし、残差分析では1位者群だけ他の群に比べて1位になりやすいことが確認された。本実験結果からも、力なき者は複数回の試験結果で判断されるよりも一発勝負に賭けるべきであると示唆された。

総合考察

本研究は、力なき者は一発勝負に賭けるべきかという問題について検討した。その結果、1回勝負は10回勝負よりも測定値に含まれる誤差が大きくなることによって、「力なき者は一発勝負に賭けるべきである」という結論が導かれた。これは1回勝負が能力の測定方法としては十分ではないために見いだされた結論と言える。

心理統計学の考え方によれば、この結論は当たり前と言えるかもしれない。廣瀬・稲垣・

深谷 (2018) は「サンプルサイズ n 個が大きくなるにつれて、得られる推定値の分散はどんどん小さくなっていきます。不偏性が成り立つ場合は、分散の値が小さい方が推定誤差は小さくなりやすいため、推定精度が高くなります。」と述べている。つまり実験2の測定値が正規分布に従うことから、10回勝負の標本分布の標準偏差は1回勝負の測定値の標準偏差よりも小さくなり（理論的には $1/\sqrt{10}$ になる）、その結果、10回勝負が課題遂行能力の異なる5群間の能力差をより正確にとらえることができたということなのである。

昨今、入学試験や様々なスポーツ大会において多様な選抜方法が用いられている。しかし、その選抜方法が正しいのかどうかを吟味することなく、我々はその結果をそのまま受け入れている。それぞれの選抜の結果は、本当に正しいのだろうか。本研究はそれぞれの選抜方法の正確さ（精度）を測定する方法を提案している。測定することができない真の値を先に確定させ、様々な試験方法や選抜方法が真の値を正確に把握しているかどうかをシミュレーション実験で確認することによって、10回勝負と1回勝負という2通りの選抜方法の有効性を評価できることを示した。スポーツ大会等ではリーグ戦、トーナメント戦、サッカーやラグビーのワールドカップのリーグ戦+トーナメント戦方式、テニス大会のシード制等々が採用されている。本研究の方法を用いることによって、それぞれの選抜方法の妥当性を評価することができると考えられる。

引用文献

- Barford, N. C. (1985) *Experimental Measurements: Precision, Error, and Truth*. John Wiley & Sons Ltd. 酒井英行 訳 (1997) 「実験精度と誤差－測定確からしさとは何か」 p.15 丸善出版
- 廣瀬雅代・稲垣佑典・深谷肇一 (2018) 「サンプリングって何だろう－統計を使って全体を知る方法」 p.21 岩波書店
- 竹綱誠一郎・稲葉凌太郎・黒住嶺 (2018) 力なき者は一発勝負賭けるべきか？—一理不尽な失敗に焦点を当てて— 人文、17号、187-195.
- 竹綱誠一郎・土井孝典・鎌原雅彦 (2016) 事前の努力量とテスト後の原因帰属との関係—努力帰属と運帰属に焦点を当てて— 人文、15号、71-82.

付記

本研究の予備実験において学習院大学大学院人文科学研究科臨床心理学専攻の石井菜々子さんにご協力いただきました。ここに記して感謝いたします。本研究は、令和5年4月に学習院大学大学院に入学した未来のサイコロジスト6名と第一著者との共同研究です。6名のサイコロジストが本研究に関わったことが6面のサイコロを実験課題に使用することにつながったことは言うまでもありません。