

心理学的性差の効果を伝えるにはどうしたらよいか 心理学を専攻しない学生を対象とした授業実践

新 美 亮 輔

第1節 問題と背景

心理学において統計的仮説検定の p 値偏重に対する批判と効果量の重要性が強調されるようになって久しい (Cumming, 2014; 大久保, 2009)。効果量を適切に解釈することの重要性は研究者などの専門家だけに関わることとみなされがちだが、実際には、心理学的知見を非専門家が理解しようとするときに起こる問題にも深く関わる。というのも、非専門家が報道やプレスリリース、一般向け読み物などを通じて心理学の研究成果に触れるとき、効果量に基づいた定量的な理解より、差の有無に基づいた定性的でカテゴリーカルな理解の方が容易かつ一般的だと考えられるからだ。たとえば文化差や性差に関する心理学的知見は、その効果量がいかに小さくても、統計的に有意な差であれば、「日本人は〇〇で、西洋人は××だ」「男は〇〇で、女は××だ」のように定性的に説明され、また理解されやすい。とりわけ民族や性別といったカテゴリーはステレオタイプの思考になじむため、こういった理解が生まれやすく、実際には小さな差が大きな差としてとらえられる可能性がある。これは、単に不正確な科学コミュニケーションとして好ましくないだけでなく、ステレオタイプや偏見を助長するおそれがあるという意味でも、問題である。ステレオタイプの特徴としてカテゴリー間差異の過大視があることは、社会心理学領域で指摘されてきた (Krueger, 1992; Sherman, Kruschke, Sherman, Percy, Petrocelli, & Conrey, 2009)。

同様の問題は大学の授業でも起こる。心理学の概論的な講義では、人間一般にあてはまる普遍的事実としてさまざまな基礎知識が講じられ、文化差や性差

に言及することは多くない。記憶や学習、ストレス反応といった基本的な心的機能について、教科書に記述できるほど大きく一貫した文化差や性差が見られることは稀なためである。しかし、特に性差は大学生が興味を持ちやすいテーマでもあり、性別による違いはないのかとか、授業中のこの説明はいわゆる「男脳」「女脳」の違いと関係があるのか、といった質問が授業コメントで出ることにはよくある。一部の精神疾患のリスクや性行動など、比較的大きな性差が実際に存在する場合もあるが、第2節でレビューするように、知覚や認知、パーソナリティといった基本的な心理特性について報告される性差はあまり大きなものではなく、男性内・女性内の個人差の方がずっと大きいことがほとんどである。とはいえ、小さくとも比較的一貫した性差が見られることもあるので、性差が全くないと説明することもできない。そこで、性差があることを授業で説明するときには、個人差に比べれば小さいこと、性差は固定的なものではなく教示や練習によって縮小したり消失したりすること（e.g., Feng, Spence, & Pratt, 2007; Steele, 2010; Tzuriel & Egozi, 2010）をあわせて説明することになる。

それでも、「性差はあるが、小さい」という説明はやはりわかりにくい。その「小ささ」を具体的にイメージできないからだ。文化差や性差が実際のところどれくらいの大きさ／小ささなのかを伝えることは、かなり難しい。心理学的特性は直接目に見えない上に、具体的に想像可能な単位のある値で表せないことが多いからだ。心理学専攻の学生に対しては、Cohen（1988）の d のような効果量の概念を導入して統計学的な説明をすることが可能だが、心理学の入門として位置づけられる概論的な科目や心理学専攻でない学生を主な対象とする教養科目では困難を伴う。同様に、日本心理学会「科学コミュニケーション・ガイドライン」（https://psych.or.jp/jpamember/press_release_guidelines/）（2021年1月21日閲覧）では研究プレスリリースでの効果量の記載が推奨されているが、研究者や専門性の高い科学記者でなければ理解が難しいだろう。

筆者は、新潟大学において2020年度に開講されたオムニバス講義「セクシュアリティ・スタディーズ」の中の1回として「心理学から見たセクシュアリティ」を担当した。この講義でテーマのひとつに性差をとりあげ、男性と女性に実際どれくらいの心理学的性差があるのか／ないのかを、心理学専攻ではな

い学生が多数を占める受講生に対し説明することにした。ここで授業実践として試みたことが2つある。第1に、心理学的性差を「あり」「なし」で説明せず、効果量の統計学的な説明なしに、性差の大きさ／小ささを伝えること。第2に、受講生がもともと心理学的性差の大きさはどの程度だと感じているのかを可視化し、実際のメタ分析研究で報告されている性差効果量と比較すること。本論文ではこの2つの試みの結果を紹介する。

第2節 心理学的性差に関するメタ分析

まず、実際の心理学研究論文で報告されている性差の効果量がどれくらいなのかレビューしておこう。性差を報告した心理学論文は数多いが、同じ研究テーマであっても、実験・調査の細かな条件設定や年齢・文化等の交絡要因、あるいは意図せぬサンプリング・バイアスにより、一貫した結果にはならないのが普通である。そのため、性差を定量的に見積もるにはメタ分析が必要になる。なお本論文では便宜的に、ある特性や課題成績について男性の方が平均値が大きいとき性差効果量 d を正の値で、女性の方が平均値が大きいとき負の値で表すことにする。

Hyde (2005, 2014) は、種々の心理学的性差のメタ分析研究を網羅的にレビューしている。学業成績から知覚・認知課題の成績、運動能力、パーソナリティ、援助行動、リーダーシップなどさまざまな領域でのメタ分析を総覧すると、見積もられた性差効果量 d は ± 0.5 を超えないものが大多数である。個々の実験・調査では性差効果量が ± 1.0 を超えるデータもしばしば見られるが、メタ分析の結果ではほとんど見られない。Hyde (2005) の Table 1 に挙げられた128特性のうち、効果量が ± 0.5 を超えるものは21で、 ± 1.0 を超えるものは心理学的特性というよりは運動能力である投擲速度 ($d = +2.18$) と投擲距離 ($d = +1.98$) の2つのみであった。こうした結果から Hyde (2005) は、男性と女性の心理学的特性はすべてではないが大部分において類似しているというジェンダー類似性仮説を提唱した。ジェンダー類似性 (gender similarities) という言葉は、性差 (gender differences/sex differences) と対になっている。性差という

言葉自体が、差の存在を暗に示唆してしまうことがあり、性差の研究とは差の存在を確証することであると考えてしまいがちなことには注意が必要であろう。これと関連して、統計的に有意な性差がなかった場合には論文として報告されにくい（出版バイアス）、メタ分析によって得られた性差効果量は実際より多少大きく見積もられている可能性も考慮する必要がある。性差の心理学研究に対するクリティカルな総説として Caplan & Caplan (2009) がある。

種々の心理学的特性の中でも、性差があると広く信じられていそうな、すなわち性ステレオタイプが存在すると考えられるものの効果量はどうかだろうか。そういった特性を授業で取り上げるにあたり、さまざまなメタ分析研究をレビューした。授業スライド上で紹介したのは次のとおりであった。心的回転 ($g = +0.57$, Maeda & Yoon, 2013; $d = +0.56$, Voyer, Voyer, & Bryden, 1995), 心的回転を含む空間認知課題 ($d = +0.37$; Voyer et al., 1995), 空間記憶課題 (-0.27 ; Voyer, Postma, Brake, & Imperato-McGinley, 2007), 言語課題 (-0.11 ; Hyde & Linn, 1988), 小学校～大学の言語科目・数学・理科 (science) の成績 (順に -0.37 , -0.07 , -0.15 ; Voyer & Voyer, 2014), 他者の非言語的感情表出 (表情など) の認識 (-0.19 ; Thompson & Voyer, 2014), パーソナリティの主要5因子のうち情緒安定性 ($+0.40$; Schmitt, Realo, Voracek, & Allik, 2008), 攻撃性 ($+0.44$; Knight, Guthrie, Page, & Fabes, 2002), 日本人の自尊感情 ($g = +0.17$; 岡田・小塩・茂垣・脇田・並川, 2015), リスクを取る行動 ($d = +0.13$ から $+0.36$; Byrnes, Miller, & Schafer, 1999; Cross, Copping, & Campbell, 2011), デートや交際における外見的魅力の重視 ($+0.47$ から $+0.54$; Feingold, 1990, Analyses 1 and 6), 笑顔の表出 (-0.41 ; LaFrance, Hecht, & Paluck, 2003), 過去の抑うつ的な反すう (-0.24 ; Johnson & Whisman, 2013), 性交渉の頻度・ゆきずりの (casual) 性交渉の頻度・マスターベーションの頻度 (順に, $+0.10$ から $+0.31$, $+0.38$ から $+0.81$, $+0.53$ から $+0.96$; Oliver & Hyde, 1993; Petersen & Hyde, 2010), 理科・数学・工学への興味 (順に, $+0.36$, $+0.34$, $+1.11$; Su, Rounds, & Armstrong, 2009)。このほか、メタ分析ではないが比較的サンプルサイズが大きいデータから求めた性差効果量として、17歳高校生の身長 ($d = +2.26$; 文部科学省2019年度学校保健統計調査による), 20-24歳の身長および握力 (順に $+2.54$, $+2.90$; スポーツ庁2019年度

体力・運動能力調査,「小都市」の結果による),単純反応時間 (-0.15 ; Deary & Der, 2005, 24歳のデータによる),日本の成人のパーソナリティの外向性 (-0.18 ; 川本・小塩・阿部・坪田・平島・伊藤・谷, 2015) も取り上げた。

心的回転課題成績, 数学や理科への興味, 攻撃性は平均的に男性の方が高く, 言語課題成績や他者の非言語的感情表出の認識は平均的に女性の方が高いなど, 一般的な性ステレオタイプの方に合致した結果も多い。しかし, 空間記憶課題や理科の成績で平均的に女性の方が高いなど, 性ステレオタイプとは逆方向の結果もあった。Hyde (2005) の結果と同様, 効果量は大部分が ± 0.5 以下であった。なお, 身長性の性差は上述のように $d = +2.0$ を超えている。身長のようなほぼ生物学的性的二型と言える性差に比べると, 心理学的性差は小さい。また, これら心理学的性差のメタ分析の結果は, 生物学的で変化しない性差の存在を意味するとは限らないことに留意する必要がある。

第3節 効果量の表現方法

これら性差効果量の数値そのものは, 効果量の概念を理解していなければ解釈できないので, そのまま授業で示すわけにはいかない。学生が性差を「あり」「なし」ではなく定量的に理解するためにはどうしたらよいのだろうか。3つの方法を試みた。

まず明らかに有用なのは, 2つの分布をグラフで示し, 効果量を視覚的に表すことである。これによって, 2つの分布のずれの程度を評価するには平均値差だけでなく分散の考慮も必要であることが容易に理解できる。たとえば大学生向けの心理学の教科書である『マイヤーズ心理学』(Myers, 2011) は, 自尊感情の性差 $d = +0.21$ を2つの正規分布を重ね描きした模式図で示し (p. 123, 図4.7), 男性内・女性内の個人差の方が性差より大きいことを説明している。そのため, 今回の講義でも図1のようにグラフを用いて効果量を図示した。なお, 男女各群の標準偏差のデータが利用可能な場合はそれも反映させて作図したが, 効果量のみがわかっている場合には等分散を仮定して作図した。

第二の方法は, 比較対象を提供することだった。やはり非専門家にとって定

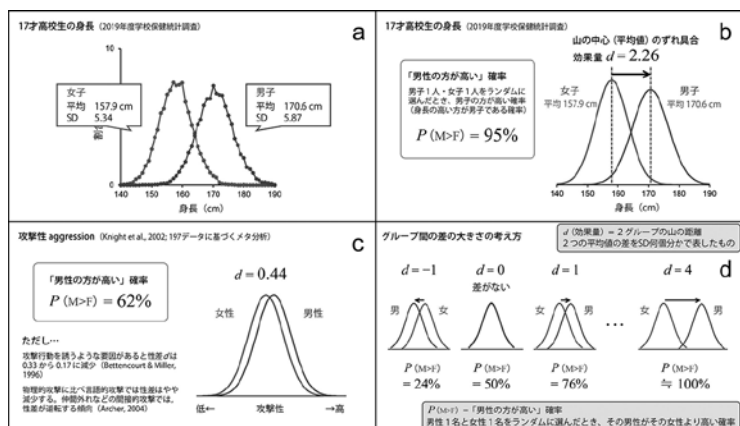


図1. 授業で用いたスライド資料の抜粋。2つの分布を重ね描きしたグラフにより、効果量 d の意味を視覚的に提示した。

量的な理解が難しい数値として、リスクがある。リスク認知や科学コミュニケーションの研究において、非常に小さな確率の値で表されることが多いリスクを非専門家が理解しやすい形で伝えるための方略がいくつか提案されているが、そのひとつに、いわゆるリスクの階梯 (risk ladder; 広田・増田・坂上, 2006, p. 232) がある。すなわち、他の身近なリスクの確率と並べて示すことで、そのリスク (たとえば災害や副作用のリスク) がどのくらい大きい／小さいのかを適切に理解しやすくなる。

そこで今回の授業では、身長を比較対象に用いた。身長のデータを授業冒頭に図示し (図1a, b), 身長の性差が $d = +2.26$ 程度であることを説明した。この際、 d の詳しい定義には触れず、単に「2つの山の離れ具合、重なっていない程度」を表す指標として、やはり図を用いて解説した (図1d)。これによって、受講生は性差の具体的なイメージを思い浮かべることができる。また、たとえば心的回転成績の性差 $d = +0.56$ (Voyer et al., 1995) について、身長の性差の4分の1程度である、と表現することができる。

身長を比較対象として導入した理由はいくつかある。第一に、実際のデータ

が官庁統計から容易に利用可能である。第二に、身長はさまざまな心理特性とは違って目に見えるものであり、しかも学生は日常的に多数の男性・女性を目にしているから、身長の性差の大きさをある程度正確に知っていると考えられる。第三に、非専門家が心の性差を考えると、「身長にはっきりした性差があるように、脳や心にも性差がある」というアナロジー的理解がなされることがある。これは、いわゆる「男脳」「女脳」の違いを強調する心理学神話・神経神話でしばしば見られる説明である。たとえばベストセラーとなった『話を聞かない男、地図が読めない女』（Pease & Pease, 1998）でも、「男女の身体は、それぞれの役割に合わせて発達していった。たいていの男は女より背が高く、力も強くなった。そして脳のほうも、役割に応じて進化していった」（p. 19）という記述が見られる。このような記述は、心理学的性差も身長性の性差と同じくらい大きなもので、しかも一生変わらないという誤ったイメージを生む。実際の心理学的性差の効果量は、第2節でレビューしたように、身長の性差よりもかなり小さい。身長と同じように心にも性差があるという誤解を解消できるという点で、身長を比較対象とすることは教育上も有効であると見込まれた。

第4節 効果量を優越確率 PS で説明する試み

効果量の定量的理解を促すために試みた第三の方法は、効果量を日常言語で理解しやすい別の指標に置き換えることである。Cohen（1988）は、2つの分布の非重複の指標（measures of nonoverlap）として、 U_1 , U_2 , U_3 の3つを提案している。いずれも効果量 d が増加すると単調に増加し、 d との換算ができる。このうち性差の指標として解釈しやすいのは U_3 であろう。いま2つの集団AとBがあり、ある特性XはBの方が平均値が高いとする。このとき、集団Bの分布のうち集団Aの平均値より高い部分の割合が U_3 である。換言すれば、集団Bからランダムに1人を選んだとき、その人の特性Xが集団AのXの平均値より高い確率である。図1a, bの身長例で言えば、17歳高校生男子のうち17歳高校生女子の平均身長より背が高い者の割合、あるいは、ある1人の17歳高校生男子の身長が17歳高校生女子の平均身長より高い確率である。この U_3 は、効果量を

d とすると、標準正規分布における $-d$ の上側確率 (d の下側確率) として与えられる。 $d=0$ のとき $U_3=50\%$ 、 $d=1$ のとき $U_3=84\%$ である。図1の身長データでは $d=+2.26$ だが、この場合 $U_3=98.8\%$ となる。心的回転の例 ($d=+0.56$) では、 $U_3=71.2\%$ となる。

Fritz, Morris, & Richler (2012) は、これら非重複の指標に加えて、優越確率 PS (probability of superiority) を紹介している。これは、上の例の集団AとBからランダムに1名ずつ選び出して特性Xを比較したときに、Aの成員よりBの成員の方がXが高い確率である。身長の例で言えば、女性と男性をランダムに1名ずつ選び出して身長を比較したとき、選び出された女性の身長より選び出された男性の身長が高い確率である。 $d=0$ のとき $PS=50\%$ である。 $d=1$ のとき $PS=76\%$ である。17歳高校生の身長の例 ($d=+2.26$) では $PS=94.5\%$ 、20–24歳の身長の例 ($d=+2.54$) では $PS=96.4\%$ となるから、任意の男性と女性を比べればたいてい男性の方が背が高いと言える。一方、心的回転の例 ($d=+0.56$) では $PS=65.4\%$ となるから、チャンスレベル (50%) を考慮すれば、任意の男性と女性を比べたときたいてい男性の方が心的回転が得意だと言うのは難しいだろう。すなわち、性別だけに基づいて、ある男性とある女性のどちらが心的回転成績が高いかを予測することは困難である。

このように PS の値は日常言語に置き換えられるため、今回の授業では U_3 ではなく PS を取り上げた。授業では、『男性の方が高い』確率 $P(M>F)$ と呼称し、性差がまったくないときは50%、すべての男性がすべての女性より高ければ100%、すべての女性がすべての男性より高ければ0%である、と説明した上で、図1dのようにグラフとともに示した。できるだけ理解を容易にするため、授業中では小数の確率ではなく百分率を用いた。なお、 PS の値はコンピュータで正規乱数を大量に発生させ (Rの `rnorm()` 関数を使用) シミュレーションを行うことによって求めた。

第5節 主観的な性差効果量を探る

性差効果量を日常言語で表現できるならば、人々が主観的に性差をどれくらい大きいととらえているのか、アンケート調査によって検討できるかもしれない。たとえば、*PS*を見積もってもらい、それがメタ分析で得られた実際の性差効果量から導かれる *PS* にどれくらい近いかを考えることができる。人々は、身長や心理学的特性の性差を現実どおりに知覚しているのだろうか、それとも過大視や過小視しているのだろうか。

この問題は、社会心理学におけるステレオタイプの正確さ (stereotype accuracy; Judd & Park, 1993; Jussim, Cain, Crawford, Harber, & Cohen, 2009) の問題に関わる。しかし、ステレオタイプはしばしばあいまいな内容を含むこともあってその正確さを評価することは難しく、またステレオタイプの正確さは個人差や題材による差も大きい (e.g., Beyer, 1999; Hall & Carter, 1999)。

ここでは、性差効果量に限定して考えよう。人々は、性差効果量のある程度正確な知識を持っているのだろうか。実は、アンケート調査で主観的な性差効果量を推定しその正確さを評価しようとした研究がある。Swim (1994) は2つの方法で性差の主観的効果量 (perceived effect size) を見積もろうとした。Study 1では、フレンドリーさ、他者の非言語的手がかりの理解能力、SAT (アメリカの大学進学適性試験) の数学の点数などの変数について、100人の男性または100人の女性が6段階の各段階にそれぞれ何人当てはまると思うかを回答するという課題が用いられた。たとえばフレンドリーさなら「非常にフレンドリーでない」から「非常にフレンドリーである」までの6段階に100人を振り分け、人数で回答する。SAT の成績なら、「200–300点」から「701–800点」の6段階である。100人の男性を想定した回答と100人の女性を想定した回答を比較して、主観的性差効果量を求めるのである。これをメタ分析研究で得られた実際の性差効果量と比較したところ、特段の性差の過大視は見られず、また主観的性差効果量と実際の性差効果量には正の相関があった。Study 2では効果量 d の値そのものを推定する課題が用いられた。参加者には、 $-.80$ 、 $-.50$ 、 $-.20$ 、 0 、 $.20$ 、 $.50$ 、 $.80$ という値が示され、負の値は順に大きな・中程度の・小さな「女性

「男性」の差、0は性差なし、正の値は順に小さな・中程度の・大きな「男性」の差を表すと説明が付された。条件によっては、いくつかの変数について実際の効果量の値の情報が与えられた。その上で、参加者はさまざまな変数について性差効果量を推定することが求められた。結果は Study 1と同様で、性差効果量の過大視は見られず、実際の効果量ともよく相関していた。Swim (1994)はこの結果から、人々が性差の大きさについて従来考えられていたよりも正確な知識を有している可能性を指摘している。しかし、ここで得られる主観的な効果量は回答方法の影響を受けることに注意が必要であろう。Swim (1994)は6段階を用いているが、これをより細かく分けたり、各レベルに付す言葉（「200-300点」など）を変えたりすれば、結果は変化しうる。そのため、人々の主観的な性差効果量の絶対的な大きさの正確さについてははっきりしたことは結論しづらいが、実際の効果量と正の相関が見られたことから、少なくとも性差効果量の相対的な序列（変数Aより変数Bの方が性差が大きい、など）については、ある程度正確な知識があると考えることができる。

第6節 優越確率 PS による主観的な性差効果量の調査

今回の授業でも、アンケートにより受講生の主観的な性差効果量を見積もる試みを行った。新型コロナウイルス感染症対策のため Zoom によるオンライン授業だったので、受講生の授業への関与を高め、またライブ感を保つため、授業前の事前アンケートと、授業中のアンケートを行った。いずれのアンケートも匿名で行われ、回答は強制ではなく、回答の有無や回答内容は授業の成績評価とは一切関連しないことが説明された上で行われた。回答の集計結果は授業中に紹介した。

事前アンケートは Google フォームを用いたオンラインアンケートだった。その内容を付録に示す。最初に回答者の性自認を「男性」「女性」「その他／回答しない」の3択で問うた。その後、性差に関する7つの質問があった。問1は日本に住む20代男性と女性の平均身長を問うた。問2から7で主観的 PS を問うた。まず問2で身長を例に PS の意味を説明し、男性1人と女性1人をランダム

ムに選び出したとき、男性が女性より身長が高い確率と、女性が男性より身長が高い確率を、合計が100%になるように推定して回答するように求めた。一方のみの回答を求めず合計100%となるような2つの値の回答を求めたのは、PSの意味の理解を助けるためであった。問3から7では、握力、心的回転、表情認識、数学への興味、数学の成績について同様の回答を求めた。

授業開始時刻後に回答が送信された7件を除外した121件の回答データを分析した。このうち、冒頭の性自認を問う質問の回答は「女性」が88件、「男性」が30件、「その他／回答しない」が3件だった。もともと授業の履修登録者が女性に偏っており、それを反映した性比となっている。

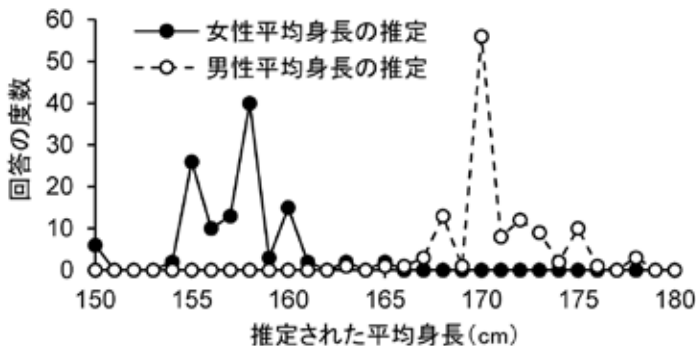


図 2. 事前アンケート問 1 (男性と女性の平均身長の推定) の回答の分布。回答者の性別はまとめてある。

問 1 (平均身長の推定) の回答分布を図 2 に示す。回答者の性別による違いは見られなかったため、全回答 ($N = 121$) をまとめて示している。女性平均身長の回答の最頻値は 158 cm、男性平均身長の回答の最頻値は 170 cm であった。女性平均身長の回答の平均は 157.1 cm, 95% CI [156.7, 157.6], 男性平均身長の回答の平均は 170.8 cm, 95% CI [170.3, 171.2] であった。これらの値は実際の日本の 20 代の女性と男性の平均身長にきわめて近く、回答者は平均身長のかかなり正確な知識を有していたと言える。自身の身長が平均と比べてどの程度か関心を

表1. 事前アンケート（主観的PS）の結果

問	主観的PS [95% CI]	対応する実際のPSと性差効果量 d
2. 身長	69.1% [66.7, 71.6]	94.5% ($d = +2.26$; 2019年度学校保健統計調査)
3. 握力	77.0% [74.9, 79.1]	98.0% ($d = +2.90$; 2019年度体力・運動能力調査)
4. 心的回転	55.4% [53.2, 57.5]	65.4% ($d = +0.56$; Voyer, Voyer, & Bryden, 1995)
5. 表情認識	38.1% [36.1, 40.1]	44.7% ($d = -0.19$; Thompson & Voyer, 2014)
6. 数学興味	61.5% [59.7, 63.4]	59.5% ($d = +0.34$; Su, Rounds, & Armstrong, 2009)
7. 数学成績	56.4% [54.7, 58.0]	48.0% ($d = -0.07$; Voyer & Voyer, 2014)

Note. PS = probability of superiority (優越確率). ここでは、ランダムに選んだ男性1名と女性1名を比較したとき男性の方が値が高い確率。

持つ大学生は多く、もともと平均身長を調べて知識を得ているケースが多いと推測される。

問2-7の回答（主観的 PS）の平均値を表1に示した。それぞれに対応する実際の性差効果量から計算された PS もあわせて示す。各問の2つの回答のうち、便宜的に1つ目の回答、すなわち男性の方が女性より高い確率の回答を分析した。なお、2つの回答の合計が100%でない回答は分析から除外した。除外された回答のいくつかは単なる暗算の誤りと思われるもの（65%と45%, など）だったが、問2-7のうち2問以上で合計が100%でなかった回答者も6名いた。問2-7の各問の有効回答数は最小で114、最大で117だった。

身長や握力の主観的 PS は実際の効果量から求めた PS よりかなり低く（表1）、この値だけを見れば、回答者は性差を過小視していたことになる。しかし、平均身長の回答は正確であった（図2）。回答者が身長の分散を過大視していたと解釈すれば整合性のある結果になるが、これは過剰な解釈と言うべきであろう。種々の認知心理学研究が示しているとおおり主観的確率は概してかなり歪むものであり（広田・増田・坂上, 2006）、平均身長の知識は正確だが性差の大きさの知識や PS の推定が不正確であってもおかしくない。心的回転（問4）でも、主観的 PS は実際の PS より50%に近く、性差が過小視されていたことになる。一方、表情判断（問5）では実際以上に女性の方が得意であると見積もられており、数値だけを見れば、ステレオタイプに沿った方向に性差が過大視

されていたことになる。数学への興味・数学の成績（問6, 7）でも、実際よりも男性の方が高いと見積もられており、ステレオタイプに沿った方向の過大視が見られた。

表1の主観的 *PS* と実際の *PS* の間には高い正の相関があった ($r = .87$)。サンプルサイズが6ときわめて小さいので参考程度ではあるが、Swim (1994) と同様の結果である。すなわち、回答者は性差効果量の相対的な序列に関するある程度正確な知識を有していたことが示唆される。特に、身長や握力といった生物学的要因の影響が強い特性に比べて心理学的性差の方が小さいという知識は（少なくとも集団レベルでは）有していたようである。

なお、回答者の性別による回答傾向の違いはあまり見られなかった。全ての問いの回答について、回答者の性自認が「男性」の場合と「女性」の場合で中心傾向が異なるかを Wilcoxon-Mann-Whitney 検定によって検討したが、有意な性差が見られたのは身長の *PS*（問2）のみで ($U = 901, p = .029$, Cliff's $d = .229$)、男性回答者の方が、男性の方が身長が高い確率をわずかに高く見積もっていた（男性の回答平均70.9%, 女性の回答平均68.4%）。他の問いでは有意な差は見られなかった ($ps > .1$)。

第7節 グラフを用いた主観的な性差効果量の推定

授業時間中にも1問のアンケートを行った。図1bのようなグラフで身長と握力のデータを示し、効果量 d の簡易な説明を行ったあと、他の心理学的性差の効果量を示す前に、性差が比較的一貫して見られる認知課題として心的回転課題を紹介してその性差効果量がどの程度だと思うかを問うた。図3枠内のスライドを示し、 $d = 0.0, 0.5, 1.0, 2.0$ の4つの選択肢の中から、実際の心的回転の性差効果量に最も近いと思うものを選んでもらった。オンラインイベントツール [slido](http://www.sli.do)(www.sli.do)のライブ投票機能を用い、投票ページのURLを画面にQRコードで表示した。受講生は匿名で回答した。

93件の回答が得られた。その結果を図3下段に示す。メタ分析で得られた実際の心的回転の性差効果量は $d = +0.56$ (Voyer et al., 1995) であり、 $d = 0.5$ の選

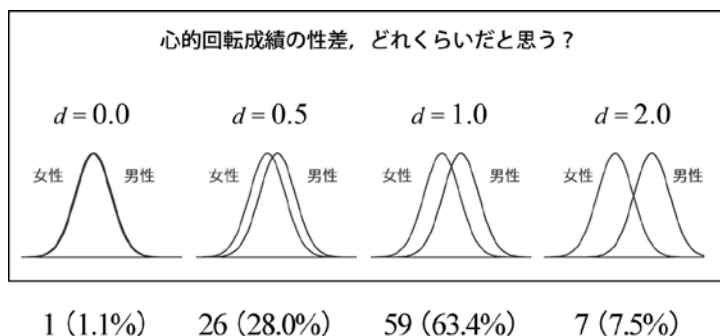


図3. 授業中アンケートの結果。心的回転の性差効果量を4択で推定してもらった。下段に各選択肢の選択数とその割合を示す。

択肢が最も近いことになるが、それより大きな $d = 1.0$ と 2.0 を選んだ回答が66件（71%）で、多くの回答者が性差を過大視していたことになる。しかし主観的 *PS* の回答ではむしろ過小視の傾向が見られており（第6節）、一貫した結果になっていない。当然のことながら、人々の回答はさまざまな知識や方略に基づいて行われており、明確な数学的関係のある複数の変数を推定してもらったとき数学的に一貫した回答になる必然性はない。したがって、これらの結果から回答者が性差効果量を実際より過大視していたか過小視していたかは、結論できないだろう。

加えて、このような質問の回答傾向は選択肢の幅に影響を受けることや、直前に身長や握力の大きな性差効果量（ $d > +2$ ）を示していたことによる係留効果（anchoring effect）の可能性も考慮する必要があるだろう。サンプリング・バイアスの可能性も指摘しておかなければならない。回答者は講義「セクシュアリティ・スタディーズ」の受講生であり、性やセクシュアリティの問題にもともと興味・関心の高い者が集まっていただろう。また、筆者の担当講義までに他の講師（文化人類学、社会学、看護学など）によるさまざまな講義を聴講しており、性差について一般的な大学生よりも知識が多い状態にあったと考えられる。実際に大学生一般が性差をどのようにとらえているかを検討するには、

研究として計画された厳密な実験・調査を行う必要がある。

第8節 まとめと展望

第1の目的、すなわち心理学的性差の大きさ／小ささをわかりやすく伝えることについて、試みた3つの方法のうち、2つの分布を重ね描きしたグラフによる図示と、比較対象（身長）を導入することは、有効であると思われた。効果量の言い換え、すなわち *PS* による表現については、性差効果量の理解に役立つかはっきりとはわからない。メタ分析で得られた実際の性差効果量から求めた *PS* の値を説明することで、性差の大きさ／小ささを日常言語で説明することには、一定の効果があるかもしれないが、今回の授業では理解度の測定などによる教育効果の定量的検証は行われなかった。今後の検討が必要である。

第2の目的、すなわち受講生の主観的な性差効果量を可視化しメタ分析による実際の性差効果量と比較することについては、困難があった。*PS* の主観的評価が実際に受講生の性差の知識を測定できているかは不明であり、したがってまた、それを実際の性差効果量と比較することにどのような意味があるかはわからない。*PS* は、実際の性差効果量を日常言語に置き換えてわかりやすく説明するためには役立つ可能性があるが、受講生の主観的効果量を可視化することに役立つかは不明である。

今回の授業では性差がテーマだったが、同様の試みは他の効果量の説明にも適応できるだろう。ただし、テーマに応じた変更や工夫が必要である。たとえば今回は比較対象として身長の性差を用いたが、もちろんこれは他のテーマでも使えるとは限らない。また、効果量を言い換える方法として今回は *PS* を用いたが、他の指標が適切なケースもあるだろう。Cohen (1988) も、非重複の指標 U_1 , U_2 , U_3 について、どれを用いるのがよいかは適用する文脈によって自由に決めるとしている。

文献

- Beyer, S. (1999). The accuracy of academic gender stereotypes. *Sex Roles*, 40, 787–813. doi: 10.1023/A:1018864803330
- Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125, 367–383.
- Caplan, P. J., & Caplan, J. B. (2009). *Thinking critically about research on sex and gender* (3rd ed.). Boston: Pearson Allyn and Bacon. (カプラン, P. J.・カプラン, J. B., 森永康子 (訳) (2010). 認知や行動に性差はあるのか 科学的研究を批判的に読み解く 北大路書房)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cross, C. P., Copping, L. T., & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin*, 137, 97–130. doi: 10.1037/a0021591
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25, 7–29. doi: 10.1177/0956797613504966
- Deary, I. J., & Der, G. (2005). Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology, and Cognition*, 12, 187–215. doi: 10.1080/13825580590969235
- Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attraction: A comparison across five research paradigms. *Journal of Personality and Social Psychology*, 59, 981–993. doi: 10.1037/0022-3514.59.5.981
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, 18, 850–855. doi: 10.1111/j.1467-9280.2007.01990.x
- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141, 2–18. doi: 10.1037/a0024338
- Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology*, 77, 350–359. doi: 10.1037/0022-3514.77.2.350
- 広田 すみれ・増田 真也・坂上 貴之 (編著) (2006). 心理学が描くリスクの世界

改訂版 慶應義塾大学出版会

- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592.
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65, 373–398. doi: 10.1146/annurev-psych-010213-115057
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53–69.
- Johnson, D. P., & Whisman, M. A. (2013). Gender differences in rumination: A meta-analysis. *Personality and Individual Differences*, 55, 367–374. doi: 10.1016/j.paid.2013.03.019
- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Bulletin*, 100, 109–128. doi: 10.1037/0033-295X.100.1.109
- Jussim, L., Cain, T. R., Crawford, J. T., Harber, K., & Cohen, F. (2009). The unbearable accuracy of stereotypes. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 199–226). New York: Psychology Press.
- 川本 哲也・小塩 真司・阿部 晋吾・坪田 祐基・平島 太郎・伊藤 大幸・谷 伊織 (2015) ビッグ・ファイブ・パーソナリティ特性の年齢差と性差：大規模横断調査による検討 発達心理学研究, 26, 107–122. doi: 10.11201/jjdp.26.107
- Knight, G. P., Guthrie, I. K., Page, M. C., & Fabes, R. A. (2002). Emotional arousal and gender differences in aggression: A meta-analysis. *Aggressive Behavior*, 28, 366–393. doi: 10.1002/ab.80011
- Krueger, J. (1992). On the overestimation of between-group differences. *European Review of Social Psychology*, 3, 31–56.
- LaFrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin*, 129, 305–334.
- Maeda, Y., & Yoon, S.-Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT:R). *Educational Psychology Review*, 25, 69–94. doi: 10.1007/s10648-012-9215-x
- Myers, D. D. (2011). *Psychology* (10th ed.). New York: Worth. (マイヤーズ, D. D., 村上 郁也 (訳). (2015). カラー版 マイヤーズ心理学 西村書店)
- 岡田 涼・小塩 真司・茂垣 まどか・脇田 貴文・並川 努 (2015). 日本人における自尊感情の性差に関するメタ分析 パーソナリティ研究, 24, 49–60. doi: 10.2132/

personality.24.49

大久保 街亜 (2009). 日本における統計改革—基礎心理学研究を資料として— 基礎心理学研究, 28, 88–93. doi: 10.14947/psychono.KJ00005878671

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin*, 114, 29–51.

Pease, A., & Pease, B. (1998). *Why men don't listen & women can't read maps: How we're different and what to do about it*. Mona Vale, Australia: Pease Training International. (ピーズ, A.・ピーズ, B., 藤井 留美 (訳) (2000). 話を聞かない男, 地図が読めない女 男脳・女脳が「謎」を解く 主婦の友社)

Petersen, J. L., & Hyde, J. S. (2010). A meta-analytic review of research on gender differences in sexuality, 1993–2007. *Psychological Bulletin*, 136, 21–38. doi: 10.1037/a0017504

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94, 168–182. doi: 10.1037/0022-3514.94.1.168

Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology*, 96, 305–323. doi: 10.1037/a0013778

Steele, C. M. (2010). *Whistling Vivaldi: How stereotypes affect us and what we can do*. New York: W. W. Norton. (ステイール, C. M., 藤原 朝子 (訳) (2020). ステレオタイプの科学 英治出版)

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, 135, 859–884. doi: 10.1037/a0017364

Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, 66, 21–36. doi: 10.1037/0022-3514.66.1.21

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion*, 28, 1164–1195. doi: 10.1080/02699931.2013.875889

Tzuriel, D., & Egozi, G. (2010). Gender differences in spatial ability of young children: The

- effects of training and processing strategies. *Child Development*, 81, 1417–1430. doi: 10.1111/j.1467-8624.2010.01482.x
- Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review*, 14, 23–38. doi: 10.3758/BF03194024
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204. doi: 10.1037/a0036620
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117, 250–270.

付録（事前アンケートの内容）

【はじめにお読みください】

- ・できるだけ講義前日（火曜）までに回答してください。
- ・1人1回だけ回答してください。
- ・検索したり他の人と相談したりせず、自分の感覚で答えてください。
- ・これはテストではありません。回答内容や回答の有無は、授業の成績評価には一切関係ありません。
- ・みなさんが男性と女性の違いをどう感じているかを調べるのが目的です。回答内容からみなさんのセクシュアリティや差別意識がわかる、ということはありません。また、回答は匿名ですので、誰がどう回答したかは教員にはわかりません。気楽に答えてください。
- ・集計結果は講義時間内に紹介する予定です。

最初に、あなたが自分の性別としてお考えのものをお答えください。

☐男性

☐女性

☐その他／回答しない

1.

日本に住む20代男性の平均身長は、何cmだと思いますか（「cm」は記入する必要はありません） _____

日本に住む20代女性の平均身長は、何cmだと思いますか（「cm」は記入する必要はありません） _____

2. 【説明をよく読んで、質問に回答してください】

日本に住む20代の中から男性1人・女性1人をランダムに選び出し、身長を比べるとします。

仮に、日本に住む20代男性の身長と日本に住む20代女性の身長に平均的な違い

がないとします。このとき、男性1人の身長が女性1人より高い確率は50%、女性1人の身長が男性1人より高い確率も50%です。

仮に、全ての男性が全ての女性より身長が高いとします。このとき、男性1人が女性1人より高い確率は100%、女性1人が男性1人より高い確率は0%です。

逆に、全ての男性が全ての女性より身長が低いとします。このとき、男性1人が女性1人より高い確率は0%、女性1人が男性1人より高い確率は100%です。

では、実際に日本に住む20代から男性1人・女性1人を選んだら、次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より身長が高い確率（%）は _____
女性1人が男性1人より身長が高い確率（%）は _____

3. 日本に住む20代の中から男性1人・女性1人をランダムに選び出し、握力を比べるとします。次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より握力が強い確率（%）は _____
女性1人が男性1人より握力が強い確率（%）は _____

4. 心的回転（mental rotation）という、空間認知能力に関する実験があります。2つの立体図形の絵を見て、3次元的に同一の立体かどうかを判断するものです。

日本に住む20代の中から男性1人・女性1人をランダムに選び出し、心的回転の実験を行ってもらったとします。次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より心的回転の成績が高い確率（%）は _____
女性1人が男性1人より心的回転の成績が高い確率（%）は _____

5. 表情判断の実験を行うとします。いろいろな表情の顔写真を見てもらい、それがどんな感情を表しているか（喜び、怒り、悲しみ、など）を当ててもらう実験です。

日本に住む20代の中から男性1人・女性1人をランダムに選び出し、この実験を行ってもらったとします。次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より表情判断の成績が高い確率（%）は _____

女性1人が男性1人より表情判断の成績が高い確率（%）は _____

6. 日本の高校生・大学生の中から、同じ学年の男性1人・女性1人をランダムに選び出し、数学にどれくらい興味があるかたずねたとします。

次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より数学への興味強い確率（%）は _____

女性1人が男性1人より数学への興味強い確率（%）は _____

7. 日本の高校生・大学生の中から、同じ学年の男性1人・女性1人をランダムに選び出し、数学の成績を比べたとします。ただし、文理は同じになるように選んだとします（つまり、2人とも文系、または2人とも理系）。

次の確率はどれくらいだと思いますか。2つの確率の合計が100%になるようにお答え下さい。

男性1人が女性1人より数学の成績が高い確率（%）は _____

女性1人が男性1人より数学の成績が高い確率（%）は _____