

СТЕРЕОТИПЫ НЕ ВЛИЯЮТ НА ЖЕНСКУЮ УСПЕШНОСТЬ В МАТЕМАТИКЕ?

Мой друг Алексей Стукальский прислал мне исследование Registered Replication Report: Johns, Schmader, & Martens (2005 [это не дата выхода работы, работа свежая]) и с огорчением заметил, что эффект угрозы подтверждения стереотипа (stereotype threat, ST) поставлен под серьезное сомнение. Это важная новость, ведь ST долгое время считалась одним из объяснений различий в результатах тестов между мужчинами и женщинами, а также между представителями разных расовых и этнических групп. В частности, широко обсуждалась идея, что девушки хуже справляются с математическими тестами именно из-за ST. Если этот фактор исключить, усиливаются позиции сторонников врожденных различий.

Давайте спокойно и честно разберемся: что на самом деле показывает новое исследование, и можно ли считать угрозу подтверждения стереотипа опровергнутой?

ВАЖНОЕ УТОЧНЕНИЕ

Дело в том, что не так-то просто найти самую актуальную версию работы, потому как в сети разошлась более ранняя, с 8 репликациями [\[1\]](#), тогда как я проанализировала чуть более позднюю – с 9. Чтобы избежать путаницы с цитатами, которые из-за разных версий не везде совпадают, я предлагаю вам либо обращаться ко мне в

личку/[комментарии](#) за полным текстом, либо ткнуть вот [сюда](#) – это должно сразу запустить скачку PDF с нужным полным текстом. Чтобы окончательно убедиться, посмотрите, что указано в Abstract: “eight direct replications” или “nine direct replications”. Нужно второе. А теперь переходим к самому разбору.

РАЗБОР

Итак, перед нами ре-анализ старой, но знаковой работы 2005 года, в которой была обнаружена угроза подтверждения стереотипа (ST). В ре-анализе же предприняли попытку перепроверить результаты, и попытка не увенчалась успехом: эффект ST подтвержден не был.

"Considering all findings of the confirmatory, preregistered analyses, we consider the results of the original study not replicated. In particular, we did not find an interaction between gender and condition on the mathematics accuracy scores. Specifically, the women in the stereotype-threat condition did not, on average, score differently from the low-threat groups for women."

Но насколько эти результаты заслуживают доверия?

СИЛЬНЫЕ СТОРОНЫ РАБОТЫ

Методологически исследование можно считать сильным. Во-первых, была пререгистрация, что уменьшает риск ложно-положительных результатов. Во-вторых, было 9 прямых репликаций эксперимента 2005 года.

"This registered replication report describes the result of nine direct replications (N = 1709) of a representative stereotypethreat study by Johns et al. (2005)"

То есть каждая лаборатория из разных стран провела свой эксперимент по максимально близкой к оригиналу методологии с общим числом участников 1709 человек. Это обеспечивает высокую внешнюю валидность.

Наконец, в-третьих, применялся верифицированный экспериментальный дизайн.

"Toni Schmader, who co-authored the original study, provided detailed feedback and help during the construction phase of this project. Moreover, she provided the original study materials (questionnaires, Graduate Record Examinations (GRE) items, advertisements, manipulations, and experimenters' protocol), and the original dataset of the study."

Это золотой стандарт репликационных исследований: Тони Шмадер (одна из авторов оригинального исследования) предоставила оригинальные материалы (инструкции,

тексты тестов, стимулы, анкеты) и оригинальный набор данных. А кроме того авторы старой работы участвовали в проверке протокола репликации.

СЛАБЫЕ СТОРОНЫ РАБОТЫ

У этой сильной работы есть и некоторые слабые стороны, которые признают сами ее авторы. Во-первых, они создали свои собственные тесты, так как опасались, что американские задания не будут корректно работать в Европе.

"Second, as we were uncertain how the items used by Johns et al. (2005) would function outside of the United States, we decided to create two new mathematics tests (see Pilot of GRE items and GRE test)."

Сам по себе это очень слабый недостаток, потому что исследователи провели пилотирование новых тестов, чтобы они были сопоставимы с оригиналом. Методология, процедуры, инструкции и анализ сохранены. Но есть кое-что, что все же усиливает данную проблему. А именно – авторы признают, что их тесты могли быть недостаточно сложными для того, чтобы эффект проявил себя.

"It may be that our mathematics tests were not difficult enough to the participants to elicit stereotype-threat effects (Keller,

2007; Spencer et al., 1999)."

Теоретически низкая сложность может приводить к эффектам, схожим с потолочными, когда различия между группами сглаживаются, однако сами авторы напрямую об этом не пишут.

Кроме того, авторы признают, что участники часто не завершали тест полностью.

"However, we do note that assessing test reliability and test difficulty is not easy considering that many participants did not finish the mathematics test (Lu & Sireci, 2007)."

Даже при высокой формальной надежности тестов ("*Our tests appeared relatively reliable ($\alpha = .82$ for the easier test and $\alpha = .88$ for the regular test*")), их реальная сложность и частая незавершенность могли ослабить проявление эффекта ST. Так что это слабое место репликации, которое не позволяет полностью исключить влияние особенностей дизайна на результат.

Пожалуй, из серьезных замечаний можно сделать только это. Более слабым, но тоже заслуживающим внимания моментом я нахожу тот факт, что выборка ограничена только студентами из "вестернизированных" стран.

"Our participant pool was large but only contained data from American and Western-European students."

ЗНАЧИТ ЭФФЕКТА ДЕЙСТВИТЕЛЬНО НЕТ?

Дело в том, что данное исследование в принципе не может расставить все точки над "и" в данном вопросе, потому что оно проверяло только мгновенное влияние: снижает ли ST результат женщин в математическом тесте здесь и сейчас, если внушение дается непосредственно перед тестом. Но остается открытым вопрос: что с более сложными и накопительными эффектами? Ведь есть немало работ, показывающих, что ST может действовать долгосрочно, например, снижая мотивацию и вовлеченность.

Вот недавнее лонгитюдное исследование [\[2\]](#) на выборке этнически миноритарных студентов в контексте STEM.

Его основные находки следующие.

"Growth analyses indicated that students' stereotype threat and perceived cost of studying science increased during college, whereas science self-efficacy, intrinsic value, and attainment value declined. Parallel growth analyses suggested that higher initial stereotype threat related to a faster decline in attainment value and faster increase in perceived costs throughout college. Higher initial levels and a steeper increase in

stereotype threat related to lower STEM GPA. Higher initial levels and a slower decline in motivation variables related to higher STEM GPA and more completed STEM courses."

Угроза стереотипа со временем усиливалась, а мотивация к изучению науки снижалась. Эти результаты показывают, что даже если эффект ST не всегда проявляется мгновенно в виде провала на тесте, он способен постепенно подтачивать мотивацию и академическую успеваемость в долгосрочной перспективе.

Эта работа тоже методологически крепкая. Исследователи использовали *parallel process growth models* — модели параллельного роста, которые позволяют одновременно отслеживать взаимосвязанные изменения угрозы стереотипа и мотивационных переменных во времени. Эти модели выявили как прямую, так и опосредованную связь угрозы с академическими исходами (STEM GPA и количество завершенных курсов).

Помимо математики, ST может негативно влиять на успешность девушек в физике. Взглянем на эту работу [\[3\]](#).

Девушки, разделяющие стереотип "мужчины лучше в физике", демонстрировали заметно худшие результаты.

"For example, in the pre-test for calculus-based physics, female students who agreed with the stereotype performed

significantly worse than female students who disagreed with the stereotype (p value = 0.003, effect size = 0.69)."

Небольшое пояснение: pre-test — это не какой-нибудь "черновой" тест, а стандартизированное тестирование до начала курса для оценки исходных знаний и навыков.

Наибольший же гендерный разрыв наблюдался именно среди студентов, согласных с этим стереотипом, с очень крупными эффектами (Cohen's $d > 1$).

"Furthermore, Table II shows that the largest differences in performance between male and female students were between male and female students who agreed with the stereotype. Three out of the four comparisons (calculus-based pre-test and post-test, algebra-based pre-test) resulted in large effect sizes (1.36, 1.02, and 1.78, respectively)."

Это данные по физике, а не математике. В этой работе не было сильного прямого внушения: в Study 1 участников лишь просили указать свой пол перед тестом, а в Study 2 оценивали их личное согласие с гендерными стереотипами. Поэтому свежий ре-анализ, который мы разбирали в начале, не может опровергнуть результаты данной работы, так как он проверял совершенно иной тип эффекта ST. Фактически, это лабораторное внушение в одном случае и внутренние убеждения в другом.

Наконец, кроме лабораторных и лонгитюдных данных, недавно появился пререгистрированный мета-анализ по фактическому воздействию stereotype threat на рабочем месте [4].

"Results showed that stereotype threat was positively related to exhaustion, identity separation, negative affect, turnover intentions, and behavioral coping, and negatively related to career aspirations, job satisfaction, organizational commitment, job engagement, job performance, positive affect, self-efficacy, and work authenticity."

Причем основные негативные эффекты на ключевые карьерные и рабочие переменные являются средними или близкими к средним по величине (не маленькими).

"Concerning career-related variables, stereotype threat was consistent with expectations negatively related to career aspirations ($\rho = -.25$) but was unrelated to career opportunities. As for job attitudes and performance, consistent with our a-priori expectations, stereotype threat was negatively related to organizational commitment ($\rho = -.30$), job satisfaction ($\rho = -.34$), job engagement ($\rho = -.29$), and job performance ($\rho = -.35$)."

Мета-анализ также показывает, что угроза стереотипа

подрывает психологическое благополучие.

“Concerning well-being, stereotype threat was positively related to exhaustion ($\rho = .31$), identity separation ($\rho = .49$), and negative affect ($\rho = .34$), and negatively related to positive affect ($\rho = -.47$), self-efficacy ($\rho = -.26$), and work authenticity ($\rho = -.36$).”

Это означает, что ST повышает эмоциональное выгорание и негативный аффект, снижает самоэффективность и ощущение аутентичности на работе — эффекты, которые нельзя игнорировать. Причем размер эффекта снова по большей части средний, а местами приближается к крупному, как видно из цитаты выше.

А в целом результаты этой работы означают, что ST работает и вне лабораторий, негативно влияя на мотивацию, вовлеченность и эффективность сотрудников в реальных социальных условиях.

НЕБОЛЬШОЕ ДОПОЛНЕНИЕ

Справедливости ради, это не первая работа, поставившая под вопрос существование острого (если дать установку прямо перед тестами) эффекта угрозы подтверждения стереотипа. Скепсис по этому поводу высказывается уже давненько, и некоторые мета-анализы не находили либо же

находили лишь маленький эффект.

ВЫВОДЫ

Свежий ре-анализ 2025 года представляет собой методологически сильную работу, практически золотой стандарт репликационного исследования. За исключением одного важного недостатка, связанного с адаптацией тестов и их возможной недостаточной сложностью, работа выглядит убедительно и, скорее всего, действительно опровергает результаты знакового исследования по ST 2005 года.

Однако речь идет лишь об узком вопросе: мгновенном эффекте ST, который внушается девушкам непосредственно перед тестами по математике. Даже будь эта работа безупречной, она не может опровергнуть другие проявления ST, например, долгосрочный накопительный негативный эффект на мотивацию. Или влияние на результаты тестов по физике при слабом внушении, проявляющееся через согласие с гендерными стереотипами (девушки, которые считают, что парни лучше в физике, действительно решают тесты хуже, чем девушки, которые так не считают).

Таким образом, Registered Replication Report 2025 года ограничивает интерпретации только ситуациями мгновенного внушения, а накопительные, корреляционные и прикладные (в реальной жизни) эффекты ST остаются подтвержденными рядом исследований. Поэтому никаких

оснований “хоронить” эффект угрозы подтверждения стереотипа в целом нет.

ССЫЛКИ

1. Stoevenbelt, A. H., Flore, P., Wicherts, J., Bergers, R., Calin-Jageman, R., Gomez, L. A., ... & Weiss, L. A. S. (2024). Registered replication report: Johns, Schmader, & Martens (2005). (https://www.researchgate.net/publication/384310302_Registered_Replication_Report_Johns_Schmader_Martens_2005)
2. Totonchi, D. A., Perez, T., Lee, Y. K., Robinson, K. A., & Linnenbrink-Garcia, L. (2021). The role of stereotype threat in ethnically minoritized students' science motivation: A four-year longitudinal study of achievement and persistence in STEM. *Contemporary educational psychology*, 67, 102015. (<https://pmc.ncbi.nlm.nih.gov/articles/PMC8775166/>)
3. Maries, A., Karim, N. I., & Singh, C. (2020). The impact of stereotype threat on gender gap in introductory physics. *arXiv preprint arXiv:2007.02311*. (https://www.researchgate.net/publication/342733083_The_impact_of_stereotype_threat_on_gender_gap_in_introduutory_physics)
4. von Hippel, C., Kühner, C., Coundouris, S. P., Lim, A., Henry, J. D., & Zacher, H. (2024). Stereotype threat at work: A meta-analysis. *Personality and Social Psychology Bulletin*, 01461672241297884. (https://www.researchgate.net/publication/387020856_Stereotype_Threat_at_Work_A_Meta-Analysis)

