

Особенности реакции индивидов на обратную связь, получаемую в условиях не связанных друг с другом заданий

Екатерина Синякова, Ашот Мирзоян

Статья поступила
в редакцию
в июле 2023 г.

Синякова Екатерина Алексеевна — магистрант, Школа экономики Эразма, Университет Эразма Роттердамского. Адрес: 50 Burgemeester Oudlaan, 3062 PA Rotterdam, Kingdom of the Netherlands. E-mail: katya.sinyakova.02@gmail.com. ORCID: <https://orcid.org/0009-0007-1603-7165> (контактное лицо для переписки)

Мирзоян Ашот Гамлетович — старший преподаватель кафедры экономики инноваций экономического факультета, Московский государственный университет им. М.В. Ломоносова. E-mail: kell56@yandex.ru. ORCID: <https://orcid.org/0009-0005-9275-0099>

Аннотация

Проведен лабораторный эксперимент, в котором приняли участие 130 студентов МГУ им. М.В. Ломоносова. Участникам предлагалось выполнить задания, не связанные друг с другом и предполагающие наличие разных навыков: тест на аналитические способности и тест на соотнесение стран с их столицами. После первого задания участникам из группы воздействия предоставлялась обратная связь: справился ли участник лучше или хуже половины остальных участников в своей группе. Такая обратная связь неинформативна в отношении нового задания, так как успех или неудача в тесте на аналитические способности не связаны с тем, знает ли индивид страны и их столицы. Установлено, что студенты реагируют на негативную обратную связь: участники из группы, получившей негативную обратную связь, значительно ниже оценивают вероятность своего успеха в следующем задании. Особенностью работы является проведение эксперимента в течение двух дней во избежание возникновения «ошибки горячей руки», описанной в предыдущих исследованиях. Для проверки гипотез используются статистические тесты и регрессионный анализ, а также бутстрэп. Получен устойчивый во времени эффект неинформативной в отношении нового задания обратной связи.

Ключевые слова

обратная связь, эксперимент, негативный сигнал, уверенность, бутстрэп

Для цитирования

Синякова Е.А., Мирзоян А.Г. (2024) Особенности реакции индивидов на обратную связь, получаемую в условиях не связанных друг с другом заданий. *Вопросы образования / Educational Studies Moscow*, no 3 (2), pp. 272–298. <https://doi.org/10.17323/vo-2024-17644>

Feedback Spillovers across Unrelated Tasks

Ekaterina Siniakova, Ashot Mirzoyan

Ekaterina A. Siniakova — Master's Student at Erasmus University Rotterdam, Erasmus School of Economics, Erasmus University Rotterdam. Address: 50 Burge-meester Oudlaan, 3062 PA Rotterdam, Kingdom of the Netherlands. E-mail: katya.sinyakova.02@gmail.com. ORCID: <https://orcid.org/0009-0007-1603-7165> (corresponding author)

Ashot G. Mirzoyan — Senior Lecturer, Department of Innovation Economics, Faculty of Economics, Lomonosov Moscow State University. E-mail: kell56@yandex.ru. ORCID: <https://orcid.org/0009-0005-9275-0099>

Abstract In this paper we describe the results of a laboratory experiment involving 130 students of Lomonosov Moscow State University. The participants were asked to perform unrelated tasks involving different skills: a test of analytical skills and a capitals quiz. After the first task, participants in the treatment group are given feedback on whether they did better or worse than half of the other participants in their group. This feedback is uninformative about the new task, as success or failure in the analytical ability test is not related to whether the individual knows the countries and their capitals. It is revealed that students respond to negative feedback: participants from the group that received negative feedback significantly lower their probability of success in the next task. The peculiarity of the work is to conduct the experiment within 2 days in order to avoid the occurrence of the "hot hand error" described in previous studies. Statistical tests and regression analysis as well as bootstrap are used to test the hypotheses. A time-stable effect of uninformative relative to the new task feedback is obtained.

Keywords feedback, experiment, asymmetric belief updating, confidence, bootstrap

For citing Siniakova E.A., Mirzoyan A.G. (2024) Feedback Spillovers across Unrelated Tasks. *Voprosy obrazovaniya / Educational Studies Moscow*, no 3 (2), pp. 272–298 (In Russian). <https://doi.org/10.17323/vo-2024-17644>

Согласно результатам недавних эмпирических исследований, обратная связь оказывает влияние как на убеждения индивидов, так и на принимаемые ими решения. Экономические агенты учитывают информацию о своих прошлых успехах, принимая решение о готовности конкурировать [Banerjee, Gupta, Villeval, 2020] или выбирая оптимальный уровень собственных усилий [Drouvelis, Paiardini, 2022].

Эффекты обратной связи представляют как теоретический, так и практический интерес. Манипулируя обратной связью, можно оказывать влияние на эффективность отдельных групп людей. Так, Х. Хермес с соавторами [Hermes et al., 2019] показали, что с помощью положительных сигналов можно мотивировать наименее успевающих школьников, стимулируя их результативность. Однако механизм реакции на обратную связь, т.е. изменение убеждений индивидов после появления новой информации, сложно описать стандартной байесовской моделью [Möbius

et al., 2022]. Восприятие обратной связи искажается из-за наличия мотивированных убеждений: доверие индивидов к тем или иным утверждениям частично основывается не на объективной информации, а на их желаниях, страхах и бессознательных мотивациях. Вследствие этого, получая положительные или отрицательные сигналы о своих способностях, экономические агенты нередко придают им разный вес [Mezulis et al., 2004] и выборочно адаптируют свои убеждения [Eil, Rao, 2011].

Экспериментально выявлены эффекты обратной связи в условиях не связанных друг с другом заданий [Huang, Murad, 2016; Banerjee, Gupta, Villeval, 2020]. Авторы предлагали участникам эксперимента выполнить несколько заданий, подобранных так, чтобы они требовали наличия разных знаний или навыков: задание на визуальное восприятие и математический тест, задание на проверку памяти и задание на меткость броска. Успех или неудача в одном из них не имеют отношения к следующему заданию и не должны учитываться при прогнозе его результатов. Тем не менее участники эксперимента учитывали информацию о прошлом успехе при выборе схемы оплаты для следующего задания. Так, участники, лучше других справившиеся с первым заданием, чаще выбирали конкурентную оплату, рассчитывая оказаться лучше других и при выполнении второго задания, и наоборот.

В качестве объяснений полученного результата рассматривают эмоциональные (улучшение настроения участников, возникновение ощущения, что это «их день») и гедонистические (намеренный поиск индивидами информации, которая могла бы повысить их самооценку) механизмы [Banerjee, Gupta, Villeval, 2020]. Еще одно возможное объяснение — «ошибка горячей руки». Впервые это понятие использовалось применительно к баскетбольным матчам: оказалось, что люди склонны преувеличивать вероятность успешного броска после серии точных попаданий [Gilovich, Vallone, Tversky, 1985]. Участники эксперимента могут руководствоваться ошибочным убеждением, что за первой победой последует вторая, а за поражением — еще одна неудача. Закономерности действия механизмов обратной связи, получаемой в условиях не связанных друг с другом заданий, нуждаются в подробном изучении.

Наше исследование призвано дать ответ на следующий вопрос: существует ли устойчивая во времени реакция индивидов на обратную связь, получаемую в условиях не связанных друг с другом заданий? Для этого мы проводим эксперимент в течение двух дней. Если влияние обратной связи сохраняется на следующий день — значит, эмоциональные факторы и «ошибку горячей руки» нельзя считать единственным объяснением эффекта, выявленного в исследованиях Л. Хуанг и З. Мурад [Huang, Murad, 2016], а также Р. Банерджи, Н.Д. Гупты и М.К. Виллевал [Banerjee, Gupta, Villeval, 2020].

Мы предлагаем новый дизайн эксперимента, который позволяет избежать возникновения «ошибки горячей руки». Для оценки эффекта воздействия мы проводим статистические тесты и используем регрессионный анализ. Для определения различий в реакции на обратную связь между юношами и девушками мы используем бутстрэп.

В результате проведенного эксперимента установлено, что участники, хуже других справившиеся с тестом на аналитические способности и информированные об этом, снижают собственную уверенность в успешности выполнения следующего задания, не связанного с первым. В то же время сигнал об успешном выполнении задания не влияет на уверенность участников: индивиды из этой группы в среднем оценивают вероятность оказаться лучше других так же, как контрольная группа, которая не получала обратной связи.

Изучение механизмов действия обратной связи актуально для сферы образования. Полученные результаты указывают на проблемы, которые могут возникнуть во время обучения: студенты необоснованно занижают оценку вероятности своего успеха, а низкая уверенность в себе оказывает негативное воздействие на эффективность обучения [Araujo de, Lagos, 2013]. Обучающиеся, недооценивающие свои способности, прилагают меньше усилий и не получают в будущем результатов, которых могли бы достичь.

1. Обзор литературы

Обратная связь оказывает сильное влияние на уверенность и производительность студентов и сотрудников компаний: в литературе описано как благоприятное [Azmat, Iriberry, 2010; Kuhnén, Tumlala, 2012; Tran, Zeckhauser, 2012; Wozniak, Harbaugh, Mayr, 2015], так и неблагоприятное [Azmat et al., 2019; Barankay, 2012; Buser, 2016; Gill, Prowse, 2012; Villeval, 2020] воздействие обратной связи. На основании результатов эмпирических исследований можно сделать вывод, что индивиды чувствительны к любому виду обратной связи: экономические агенты реагируют и на зашумленную (*noisy*), и на недостоверную информацию [Lovász et al., 2022; Möbius et al., 2022].

Изучение механизмов обратной связи затрудняется рядом факторов, воздействующих на восприятие информации экономическими агентами. Наличие у индивида скрытых желаний и мотиваций может привести к некорректной оценке собственных способностей и отказу доверять негативной информации. Три ключевые особенности мотивированных убеждений отличают их от когнитивных искажений: экономические агенты искажают свои убеждения намеренно, ради получения выгоды, (*endogenous directionality*), эти искажения основаны преимущественно на эмоциях (*heat versus light*) и не описываются теорией ограниченной

рациональности (*neither naiveté nor lack of attention*) [Bénabou, Tirole, 2016]. Так, люди склонны приписывать свои успехи собственным способностям, а неудачи — невезению [Miller, Ross, 1975]. Руководствуясь своими убеждениями, экономические агенты могут избегать получения негативной информации [Burks et al., 2013; Carrillo, Mariotti, 2000; Karlsson, Loewenstein, Seppi, 2009; Sweeny et al., 2010], намеренно избегать воспоминаний о тех или иных событиях [Zimmermann, 2020] или неверно трактовать полученные сигналы, принимая свои желания за действительное положение дел [Mayraz, 2011].

Исследователи, изучающие возникающую асимметрию — различное влияние положительной и негативной обратной связи, — не достигли консенсуса относительно ее причин. Так, в одних работах показано, что люди не принимают во внимание плохие новости [Moore, Healy, 2008; Sharot, 2011], наделяют большим весом положительную информацию о себе [Eil, Rao, 2011] или рассматривают позитивные сигналы как более информативные [Möbius et al., 2022]. В других исследованиях таких различий в восприятии и оценке позитивной и негативной обратной связи не обнаружено [Barron, 2021; Grossman, Owens, 2012; Shah et al., 2016] или же получен обратный результат: выяснилось, что экономические агенты больше доверяют негативным сигналам [Ertac, 2011]. Более того, индивиды склонны к «консерватизму»: после получения информации о результатах своих действий участники экспериментов корректируют убеждения о собственных способностях в меньшей степени, чем можно было предполагать исходя из байесовской модели принятия решений [Möbius et al., 2022].

В реакции на обратную связь наблюдаются гендерные различия. Так, при получении зашумленной негативной обратной связи юноши объективно оценивают степень влияния случайных факторов, в то время как девушки винят себя в недостаточном уровне усилий или нехватке способностей [Shastri, Shurchkov, Xia, 2020]. В условиях конкуренции женщины с меньшей вероятностью принимают решение продолжать участие в соревновании после проигрыша на ранних этапах [Buser, Yuan, 2016], у мужчин же после неудачи значительно увеличивается неприятие риска [Filippin, Gioia, 2018].

Обратная связь в условиях заданий, не связанных друг с другом и предполагающих наличие разных способностей, — одна из наименее исследованных разновидностей обратной связи. Уже упоминавшееся исследование Л. Хуанг и З. Мурад [Huang, Murad, 2016] было построено следующим образом: участникам предлагали задание на визуальное восприятие и тест на уровень математических способностей. Первое задание не связано со вторым. Более того, оно содержит элемент случайности: участникам нужно за короткое время определить, на какой картинке (справа или слева) изображено больше точек, и в заданиях повышено-

го уровня сложности картинки различались только на одну точку. Тем не менее участники, справившиеся с заданием лучше других, оценивали уверенность в себе в задании на уровень математических способностей выше, и наоборот. Р. Банерджи, Н.Д. Гупта и М.К. Виллевал [Banerjee, Gupta, Villeval, 2020] предложили дизайн эксперимента, состоящий из теста на проверку памяти и задания на точность броска: участникам нужно было забросить 15 мячей в стоящее на расстоянии 3,5 метра от них ведро. Авторы показали, что несмотря на отсутствие связи между заданиями, участники, получившие позитивный сигнал после первого задания, были больше уверены в своей победе во втором задании, и наоборот.

Основной вклад нашей работы заключается в выявлении устойчивого во времени эффекта обратной связи, получаемой в условиях не связанных друг с другом заданий. По результатам проведения эксперимента мы показали, что участники, хуже справившиеся с заданием в первый день и получившие информацию об этом, во второй день значительно ниже оценивали свои шансы на успех в последнем задании.

2. Гипотезы В работе проверяются следующие гипотезы.

H1. Обратная связь по результатам первого задания влияет на уверенность участников в своем успехе при выполнении следующего задания: индивиды, лучше других справившиеся с тестом на аналитические способности, будут более уверены в своем успехе при выполнении теста на знание фактов, хотя эти задания не связаны друг с другом. И наоборот, индивиды, хуже других справившиеся с тестом на аналитические способности, будут менее уверены в своем успехе при выполнении теста на знание фактов.

Гипотеза основана на результатах, полученных Л. Хуанг и З. Мурад [Huang, Murad, 2016] и Р. Банерджи, Н.Д. Гуптой и М.К. Виллевал [Banerjee, Gupta, Villeval, 2020]. В данной работе мы ожидаем, что результаты теста на аналитические способности отразятся на оценке участниками вероятности своего успеха в задании на соотнесение стран с их столицами.

H2. Обратная связь по результатам выполнения первого задания влияет на будущее решение участников об уровне прилагаемых усилий.

Как Л. Хуанг и З. Мурад [Huang, Murad, 2016], так и Р. Банерджи, Н.Д. Гупта и М.К. Виллевал [Banerjee, Gupta, Villeval, 2020] оценивали влияние обратной связи на желание участников кон-

куруировать. Авторы показали, что индивиды, получившие по результатам первого задания информацию о том, что они справились лучше медианного участника, в среднем чаще предпочитали конкурентную оплату сдельной во втором задании. В своем исследовании мы также предполагаем, что обратная связь влияет на принимаемые участниками решения, но оцениваем не желание конкурировать, а уровень прилагаемых участниками усилий. М. Друвелис и П. Пайардини [Drouvelis, Paiardini, 2022] утверждают, что участники выбирают уровень усилий в зависимости от полученной информации о своих успехах относительно соперников.

Н3. Существуют гендерные различия в реакции на обратную связь.

Согласно ряду исследований [Buser, Yuan, 2016; Filippin, Gioia, 2018], юноши и девушки по-разному реагируют на обратную связь. С одной стороны, у девушек самооценка изначально ниже, чем у юношей, и они склонны преуменьшать свои достижения [Steegeh et al., 2021; Verna, 2002] и винить себя в том, что находится за пределами их контроля, например в бесплодии [McLeod, Ponnese, 2008]. С другой стороны, некоторые исследователи [Brandts, Groenert, Rott, 2015; Wozniak, Harbaugh, Mayr, 2014] утверждают, что юноши сильнее корректируют уверенность в себе после получения обратной связи о собственных способностях. Мы сравниваем воздействие обратной связи, полученной в условиях несвязанных заданий, на оценку юношами и девушками вероятности собственного успеха, а также на выбираемый ими уровень усилий.

Н4. Участники, выражающие бóльшую степень согласия с утверждениями о связи между успехами и неудачами в разных сферах жизни, сильнее реагируют на обратную связь.

Р. Банерджи, Н.Д. Гупта и М.К. Виллевал [Banerjee, Gupta, Villeval, 2020] высказали предположение, что наличие реакции на обратную связь в рамках не связанных друг с другом заданий, проверяющих разные навыки, объясняется, в частности, когнитивными искажениями и ошибочной убежденностью участников эксперимента в наличии взаимосвязи между этими задачами. Л.В. Линарес, А.Р.Р. Торрес и С.Р. Перейра [Linhares, Torres, Pereira, 2022] показывают, что по степени согласия с рядом специально подобранных утверждений можно судить, насколько индивид верит в «справедливый мир». Мы предполагаем, что участники, выражающие согласие с утверждениями о связи между успехами и неудачами в разных сферах, в большей степени ошибочно убеждены в наличии связи между разными заданиями и сильнее реагируют на обратную связь.

3. Данные В проведенном нами лабораторном эксперименте приняли участие 130 студентов 1–4-го годов обучения на бакалаврских программах экономического, биологического, филологического, юридического и геологического факультетов Московского государственного университета им. М.В. Ломоносова.

Эксперимент создан на платформе *oTree* и размещен на сервере *Heroku*. Участники выполняли задания в классах, оснащенных компьютерами, переходя по индивидуальным ссылкам.

Исследование проходило в два этапа. Для тестирования дизайна эксперимента сначала организовано пилотное испытание с 39 участниками и интервью после выполнения всех заданий. Основную выборку, на которой проверяются гипотезы, составил 91 человек. Для шести участников отсутствует информация по одной из переменных (оценка уверенности в себе), поэтому при проверке части гипотез используются 85 наблюдений.

Задания эксперимента следовало выполнять на протяжении двух дней. В первый день обязательным условием было совместное начало и окончание эксперимента группой участников, приглашенных в одно время (от 5 до 20 человек). Во второй день каждый участник мог подбирать удобное время начала эксперимента и самостоятельно принимать решение о его окончании. Прохождение эксперимента в среднем занимало от 30 до 40 минут каждый день. Все участники, пришедшие в первый день, присутствовали также и во второй день проведения эксперимента и прошли все этапы.

Все ответы обезличены. Для объединения результатов, полученных за два дня, участники придумывали индивидуальный идентификатор, который они указывали в начале каждого дня испытаний. Описательная статистика полученных данных представлена в Приложении 1.

4. Дизайн эксперимента
4.1. Распределение по группам

В первый день эксперимента перед началом выполнения заданий проводилась рандомизация: каждый участник случайным образом распределялся в контрольную группу или группу воздействия. Такой способ рандомизации обеспечивает внутреннюю валидность эксперимента: несмотря на то что исследование проводилось на протяжении нескольких дней, группы контроля и воздействия однородны, а внешние факторы, если они возникают в тот или иной день проведения эксперимента, одинаково отражаются на группе воздействия и группе контроля.

По окончании первого дня эксперимента каждый участник из группы воздействия получал бинарный сигнал: оказался ли его результат выше или ниже медианного результата среди тех, кто проходил эксперимент в одно время с ним. Информацию о попадании в ту половину участников, которая лучше справилась с испытанием, далее будем называть позитивным сигналом. Если

же студенту из группы воздействия сообщали, что его результат оказался ниже медианного, будем считать, что он получил негативный сигнал.

Таким образом, все наблюдения можно разделить на три группы: получившие позитивный сигнал, получившие негативный сигнал и группа контроля. Распределение участников по группам представлено в табл. 1. Из-за небольшого числа наблюдений медианное значение нередко встречалось сразу у нескольких участников, все они были отнесены в одну из групп, поэтому число участников, получивших позитивный и негативный сигнал, оказалось неодинаковым.

Таблица 1. Число участников, попавших в контрольную группу и группы воздействия

Обратная связь	Всего	Девушки	Юноши
Без обратной связи	29	20	9
Негативный сигнал	30	16	14
Позитивный сигнал	32	24	8
Всего 91 человек			

4.2. Общая концепция

На рис. 1 представлена схема эксперимента. Участникам предлагалось выполнить три задания: тест на аналитические способности, задание на измерение уровня усилий, тест на соотнесение стран с их столицами. Они не получали вознаграждения за факт прохождения испытания, но имели возможность заработать до 1000 рублей по результатам выполнения заданий. В первый день перед началом выполнения задания все студенты заполняли небольшую анкету, указывая свои демографические характеристики (пол, возраст), а также определяя степень согласия (от 0 до 100%) с двумя утверждениями о связи между успехами и неудачами в разных сферах жизни.

Утверждение 1. Оступившись один раз, легче оступиться вновь.

Утверждение 2. Талантливый человек талантлив во всем.

Далее участников просили оценить вероятность, с которой они рассчитывают оказаться лучше половины своих соперников, т.е. половины студентов, пришедших на эксперимент в одно время с ними, в задании на уровень аналитических способностей. После этого участники переходили к выполнению первого задания. Окончившие работу ждали, пока все остальные выполняют задание. Как только последний участник завершал прохождение теста, все автоматически переключались на заключительную страницу. Те, кто попал в группу воздействия, видели сообщение о том, оказался ли их результат выше или ниже медианного значения.

Рис. 1. Общая схема дизайна эксперимента

Вознаграждение за факт участия:
0 рублей



В начале второго дня эксперимента участникам сообщали, что испытание в этот день не будет связано с первым: подчеркивалось, что новое задание направлено на проверку знания фактов, а не на определение уровня логического мышления. Далее каждый должен был вновь оценить вероятность, с которой он рассчитывает оказаться лучше половины соперников (тех же, что и в первый день) по результатам нового задания. После этого участникам предлагалось выполнить задание за вознаграждение. Задача этой части эксперимента состояла в измерении уровня усилий, которые готовы приложить участники для достижения успеха, и оценке его зависимости от уверенности индивида в собственном успехе. С этой целью участникам сообщали, что до начала тестирования они могут заработать, перепечатавая строки с бумажного носителя: за каждую правильно перепечатанную строку (максимум 50 строк) можно получить 20 рублей, но вознаграждение будет выплачено только тем участникам, кто впоследствии окажется среди 20% лучших в своей группе в задании на соотношение стран с их столицами. Участники самостоятельно принимали решение о количестве перепечатываемых строк и в любой момент могли завершить эту работу и сразу перейти к заключительному заданию. Наконец, последняя часть эксперимента представляла собой задание на соотношение стран с их столицами.

4.3. Задания: первый день

В качестве задания на уровень аналитических способностей использован тест Равена, разработанный в 1936 г. для оценки уровня интеллекта [Raven, Raven, Court, 1936]. Он направлен на проверку аналитического мышления, что предположительно оказывает сильное влияние на самооценку участников, а потому и на их уверенность в собственных способностях [Möbius et al., 2022]. Кроме того, задание нерутинное и поэтому может заставить участников

прилагать больше усилий и относиться к результату как к оценке собственных способностей. Участникам предлагалось выполнить как можно больше заданий (максимум 20) в течение 10 минут (рис. П1 в Приложении 2). Для удобства участников задания размещались по пять на одной странице. На первые две страницы отводилось по 125 секунд, на третью и четвертую — по 175 секунд.

4.4. Задания: второй день

Во второй день прохождения исследования участникам предлагались два задания: задание для измерения уровня усилий и тест на соотнесение стран и столиц. Задание для измерения уровня усилий состояло из 50 раундов, и в любой момент участник мог пропустить задание и перейти к заключительному тесту (участники были предупреждены о том, что после теста эксперимент будет окончен).

В пилотной версии исследования, в которой участвовали 39 студентов, задание на уровень усилий было следующим: в каждом раунде на экране появлялся короткий набор латинских букв, и требовалось указать, сколько раз в нем встречается определенная буква. Это задание в среднем не занимало более 15 минут. В ходе проведенных после эксперимента интервью стало ясно, что участникам задание показалось интересным, прежде всего за счет необычного шрифта и быстрой смены раундов, так что значительная часть участников выполнила все 50 заданий.

Перед проведением основного эксперимента было принято решение изменить дизайн, усложнив задание для измерения уровня усилий. Участники получали лист А4 с 50 строками несвязного набора латинских букв и цифр. В каждый из 50 раундов участникам предлагалось перепечатать строку с указанным на экране номером (не по порядку), пропуская определенную букву (рис. П2 в Приложении 2).

Последнее задание, тест на знание фактов, заключалось в соотнесении 15 стран с их столицами (рис. П3 в Приложении 2). Во избежание списывания время выполнения задания ограничивалось 75 секундами: иначе, несмотря на присутствие организатора в аудитории, участник мог найти список стран и столиц мира и указать верные ответы, пока организатор отвечает на вопрос другого участника.

5. Методология и результаты

В табл. 2 указаны средние значения показателей уверенности участников в собственном успехе и средние результаты выполнения всех заданий отдельно для каждой экспериментальной группы. Выборка сбалансирована по априорной оценке участниками вероятности своего успеха: все группы в среднем одинаково определяют вероятность оказаться лучше половины соперников в

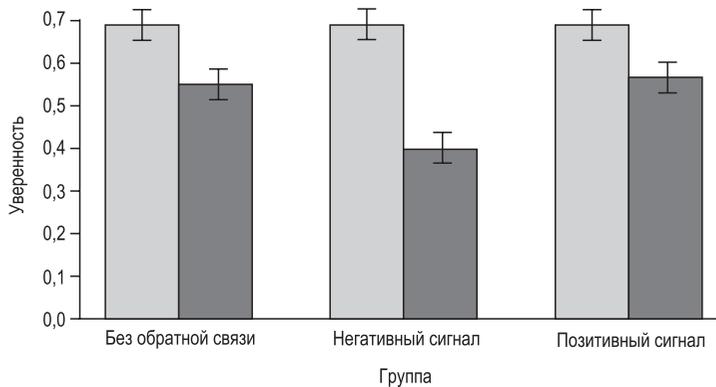
первый день (p -value для t -критерия $> 0,1$). Среднее значение числа верно указанных столиц статистически не различимо по группам (p -value для t -критерия $> 0,1$): студенты, успешные и менее успешные в тесте на аналитические способности и получившие, соответственно, разную обратную связь, в среднем одинаково справляются с заданием на соотнесение стран с их столицами, как и предполагалось в дизайне эксперимента.

Таблица 2. Средние значения по группам, сформированным по признаку получения обратной связи

Обратная связь	Уверенность в первый день	Уверенность во второй день	Верные ответы	Столицы	Матрицы Равена
Без обратной связи	0,7	0,5	23,0	6,1	9,0
Негативный сигнал	0,7	0,4	19,5	5,4	6,3
Позитивный сигнал	0,7	0,6	24,0	5,9	11,2

Для проверки первой гипотезы проведен тест на равенство средних значений оценки вероятности оказаться среди лучшей половины участников в каждой экспериментальной группе. На рис. 2 приведены средние значения для разных групп в первый и во второй день эксперимента с 90%-ными доверительными интервалами. В первый день средняя уверенность участников в собственном успехе была одинаковой для всех групп. Во второй день средняя оценка вероятности справиться лучше своих соперников снизилась во всех группах (для групп без обратной связи и с положительной обратной связью p -value $< 0,1$; для группы с негативной обратной связью p -value $< 0,05$). Общее снижение оценки вероятности превзойти остальных может быть вызвано тем,

Рис. 2. Средние значения оценки вероятности оказаться лучше половины соперников с 90%-ными доверительными интервалами



Примечание: Светло-серым цветом обозначен первый день эксперимента, темно-серым – второй день эксперимента.

что участники в среднем плохо справились с заданием в первый день. Медианный результат составил 9 решенных матриц из 20 (табл. П1 в Приложении 1). Другое возможное объяснение состоит в том, что участники в среднем действительно не уверены в своих знаниях фактов (о конкретном содержании задания участникам не сообщалось).

Тем не менее уверенность в собственном успехе во второй день эксперимента у участников, получивших негативный сигнал, значимо ниже, чем в контрольной группе (p -value = 0,08, статистическая мощность теста = 0,47). В то же время контрольная группа статистически неразличима с группой участников, получивших положительную обратную связь (p -value = 0,83).

Для проверки гипотезы об изменении уверенности участников в собственном успехе построена также множественная регрессия (табл. 3) с использованием переменной, отвечающей за получение обратной связи, и контрольными переменными априорной вероятности и пола. Коэффициент перед бинарной переменной для негативного сигнала отрицательный и значим на 5%-ном уровне. Таким образом, участники реагируют на негативный сигнал, снижая свою уверенность в успехе, несмотря на его неинформативность в отношении нового испытания.

Таблица 3. Оценка изменения уверенности во второй день эксперимента

	(1)
Константа	-0,058 (0,049)
Негативный сигнал	-0,147* (0,077)
Позитивный сигнал	-0,003 (0,061)
Пол (Женский)	-0,116** (0,057)
Количество наблюдений	85
R^2	0,37
Скорректированный R^2	0,34
F -статистика	2,55***

Примечание: Робастные стандартные ошибки в скобках. Уровни значимости: * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.

Источник: Расчеты авторов.

Участники изначально были склонны переоценивать собственные способности: в первый день среднее значение оценки вероятности оказаться лучше половины соперников составляет 0,69 (табл. П1 в Приложении 1). Такие же данные получены У. Свен-

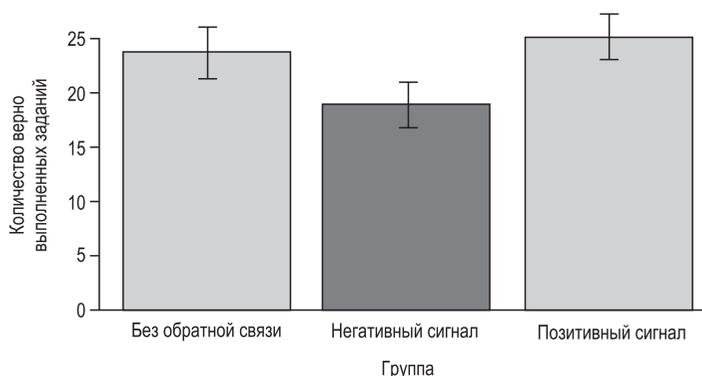
соном [Svenson, 1981]. Однако в реакции на обратную связь наблюдается асимметрия в пользу негативного сигнала, что не соответствует выводам ряда исследований [Buser, Gerhards, van der Weele, 2018; Eil, Rao, 2011; Möbius et al., 2022; Schwardmann, van der Weele, 2019; Sharot, 2011]. Такой результат, возможно, обусловлен особенностями среды обучения студентов, связанными с балльно-рейтинговой системой, такими как жесткость дедлайнов, критерии оценивания письменных работ, наличие результатов письменных работ в общем доступе. Для достижения внешней валидности необходима репликация эксперимента на выборке, состоящей из студентов высших учебных заведений со схожей образовательной средой.

Для проверки второй гипотезы — о влиянии обратной связи на уровень прилагаемых усилий — проведен тест на разницу средних для количества корректно перепечатанных строк в задании с денежным вознаграждением в каждой экспериментальной группе. Оценивались правильно перепечатанные строки, а не общее количество выполненных заданий (из 50), так как оно не в полной мере отражает уровень прилагаемых участниками усилий: некоторые участники проходили практически все раунды задания с вознаграждением, но выполняли его быстро и неаккуратно, представляя неверные ответы.

Статистически значимого различия между экспериментальными группами не обнаружено: *p-value* для теста на равенство средних составило 0,15. В связи с небольшим количеством наблюдений проведенные тесты обладают низкой статистической мощностью. Так, для теста на равенство средних в группе контроля и группе, получившей негативный сигнал, мощность составила 0,28. Для получения статистически значимого результата необходимо наличие не менее 161 человека в каждой группе. Разница в средних между группой контроля и группой, получившей положительную обратную связь, составила всего 0,15 стандартного отклонения, что не позволило бы говорить о наличии эффекта воздействия даже на достаточно большой выборке из нескольких сотен участников (рис. 3).

Отсутствие значимых различий между группами может объясняться тем, что дизайн оказался не чувствителен к различиям в уверенности участников. Большую часть выборки составили участники, которые либо выполняли все задания, либо не выполняли ни одного, в зависимости от своей мотивации. Участники, для которых потенциальное вознаграждение в 1000 рублей было существенным стимулом, старались выполнить как можно больше заданий, вне зависимости от оценки вероятности своего успеха. Участники, не заинтересованные в денежном вознаграждении, отказывались тратить время и сразу приступали к выполнению теста на знание фактов.

Рис. 3. Средние значения с 90%-ными доверительными интервалами для количества верно перепечатанных строк



Примечание: Светло-серым цветом обозначен первый день эксперимента, темно-серым – второй день эксперимента.

Для изучения связи между уверенностью участников в собственном успехе и выбираемым ими уровнем усилий также построена множественная регрессия, где в качестве зависимой переменной выступает количество ответов в задании на соотношение столиц (табл. 4). Положительный и значимый коэффициент перед уверенностью в собственном успехе может косвенно свидетельствовать о том, что участники выбирают уровень усилий в зависимости от оценки вероятности собственного успеха. Так, участники, выше оценивающие вероятность своей победы, стараются вспомнить как можно больше правильных ответов и пройти тест полностью. Участники, менее уверенные в себе, не стремятся успеть ответить на все вопросы и в ряде заданий не выбирают ни одного из предложенных вариантов ответа. Таким образом, вторая гипотеза частично подтверждается.

Таблица 4. Оценка связи между уверенностью участников и количеством указанных ответов в задании на соотношение стран с их столицами

	Количество указанных столиц
Константа	9,569*** (0,991)
Уверенность во второй день	2,530** (1,165)
Пол (Женский)	0,194 (0,677)
Количество наблюдений	85
R^2	0,43
Скорректированный R^2	0,4

Примечание: Робастные стандартные ошибки в скобках. Уровни значимости: * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.

Источник: Расчеты авторов.

Для проверки третьей гипотезы — о гендерных различиях — мы разделили каждую группу на две по признаку пола. В связи с наибольшим количеством наблюдений использован бутстрэп.

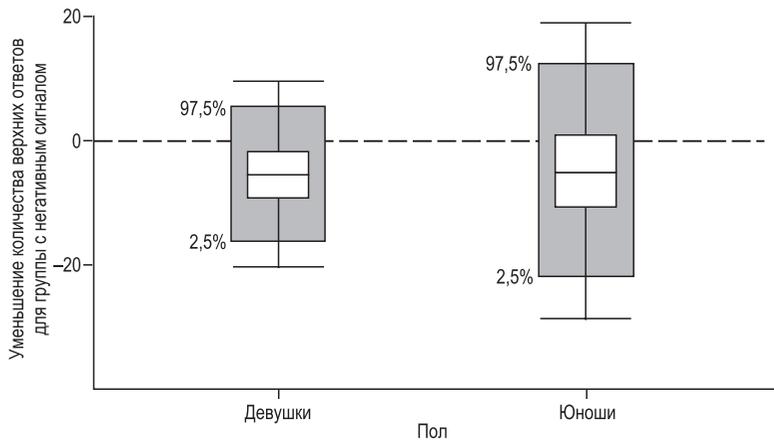
Для оценки гендерных различий в изменении уровня прилагаемых усилий отдельно оценивается регрессия на подвыборке, состоящей только из юношей, и подвыборке, состоящей только из девушек. Для каждой регрессии строится доверительный интервал оценки коэффициента β_1 в уравнении (1):

$$\text{Количество верных ответов}_i = \beta_0 + \beta_1 \cdot \text{Негативный сигнал}_i + \beta_2 \cdot \text{Позитивный сигнал}_i + \varepsilon_i, \quad (1)$$

где ε_i — случайная величина, имеющая нормальное распределение и нулевое математическое ожидание.

Мы используем алгоритм, описанный в [Davison, Hinkley, 1997], предполагая отсутствие гетероскедастичности: *p-value* в тесте Бреуша — Пагана составило 0,3 (Приложение 3, алгоритм). На рис. 4 построен ящик с усами для бутстрэпируемой оценки.

Рис. 4. Ящики с усами для оценок, полученных с помощью бутстрэпа. Серым цветом обозначены 95%-ные доверительные интервалы



Поскольку использование перцентильного доверительного интервала, указанного на рис. 4, не рекомендуется для оценки статистик, чувствительных к выбросам [Wilcox, Keselman, 2003], мы отдельно строим студентизированный доверительный интервал (формула приведена в Приложении 3).

Студентизированный доверительный интервал для оценки коэффициента перед переменной, отвечающей за получение негативного сигнала, для юношей: $[-22,25; 12,73]$, для девушек: $[-16,23; 5,50]$. Нуль входит в доверительный интервал в обоих случаях, коэффициент не значим. Следовательно, негативный сигнал не оказывает влияния на принимаемое участниками решение об

уровне прилагаемых усилий в задании с денежным вознаграждением.

Для проверки гипотезы о наличии гендерных различий в изменении уверенности в себе на отдельных выборках для юношей и для девушек строится доверительный интервал для оценки коэффициента β_2 в уравнении (2):

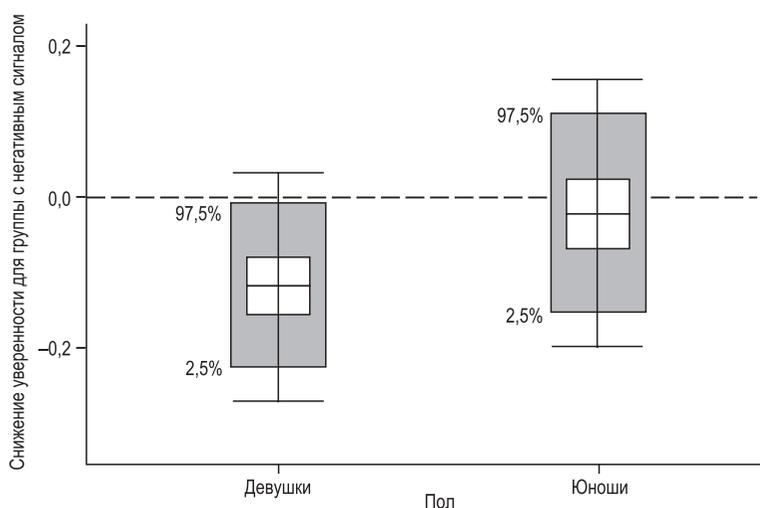
$$\begin{aligned} \ln(\text{уверенность во второй день})_i &= \\ &= \ln(\text{уверенность в первый день})_i + \beta_2 \text{ негативный сигнал}_i + \\ &+ \beta_3 \cdot \text{позитивный сигнал}_i + \varepsilon_i, \end{aligned} \quad (2)$$

где ε_i — случайная величина, имеющая нормальное распределение с математическим ожиданием, равным нулю.

В модели (2), построенной на исходной выборке, присутствует гетероскедастичность (p -value для теста Бреуша — Пагана составляет 0,001) — а значит, использование алгоритма, приведенного в Приложении 3, некорректно: дисперсия остатков зависит от значений регрессоров. В таких случаях используется механизм дикого (*wild*) бутстрэпа [Wu, 2007], при котором скорректированные остатки модели r_i^{new} соответствуют наблюдениям с тем же индексом. В данной работе используется алгоритм, предложенный в [Mammen, 1992].

Распределение бутстрэпируемой статистики с 95%-ным доверительным интервалом показано на рис. 5.

Рис. 5. Ящики с усами для оценок, полученных с помощью бутстрэпа. Серым цветом обозначены 95%-ные доверительные интервалы



Стьюдентизированный доверительный интервал для оценки коэффициента при переменной, отвечающей за получение негативного сигнала, для юношей $[-0,12; 0,11]$ включает нуль, а для

девушек $[-0,25; -0,05]$ — нет. Таким образом, эффект от негативного сигнала проявляется только у девушек, а юноши не реагируют на обратную связь. Следует отметить, что девушки в среднем хуже юношей справились с заданием на соотнесение стран с их столицами, в то время как в тесте Равена значимых различий не выявлено (табл. П2 в Приложении 1). Различия в баллах между юношами и девушками во второй день могут отчасти объясняться разным уровнем уверенности в успехе, который до некоторой степени обусловлен неодинаковой реакцией на обратную связь. Третья гипотеза подтвердилась.

Для проверки четвертой гипотезы — о связи убеждений участников с их реакцией на обратную связь — построена регрессия (табл. 5) с добавлением произведения переменной, отвечающей за получение обратной связи, и бинарной переменной, отвечающей за степень согласия с утверждениями: «Оступившись один раз, легче оступиться вновь» (утверждение 1) и «Талантливый человек талантлив во всем» (утверждение 2).

Таблица 5. Оценка уверенности в вероятности собственного успеха во второй день эксперимента. Проверка гипотезы 4

	Логарифм уверенности во второй день
Константа	0,114** (0,054)
Логарифм уверенности в первый день	0,595*** (0,097)
Негативный сигнал	-0,120** (0,060)
Позитивный сигнал	-0,005 (0,047)
Негативный сигнал*Степень согласия с утверждениями 1 и 2	0,043 (0,085)
Позитивный сигнал*Степень согласия с утверждениями 1 и 2	0,058 (0,052)
Количество наблюдений	85
R^2	0,37
Скорректированный R^2	0,34

Примечание: Робастные стандартные ошибки в скобках. Уровни значимости: * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.

Источник: Расчеты авторов.

Переменная принимает значение, равное единице, для тех участников, которые согласны с каждым из утверждений не менее чем на 30%, и равно нулю — для всех остальных. Мы рассматри-

вали разные способы создания переменной, изменяя пороговое значение степени согласия (табл. ПЗ в Приложении 4). Результаты устойчивы по отношению к выбору критерия. Четвертая гипотеза не принимается, так как оба коэффициента перед произведениями — для позитивного и негативного сигналов — не значимы.

6. Заключение В статье приведены результаты лабораторного эксперимента, в котором приняли участие 130 студентов, обучающихся на бакалаврских программах экономического, биологического, филологического, юридического и геологического факультетов Московского государственного университета им. М.В. Ломоносова.

Участникам было предложено выполнить задания, не связанные друг с другом и предполагающие наличие разных навыков: тест на аналитические способности и тест на соотнесение стран с их столицами. По результатам первого задания участникам из группы воздействия предоставлялся бинарный сигнал об их успешности: справились ли они лучше (позитивный сигнал) или хуже (негативный сигнал) половины участников, проходивших эксперимент в одно время с ними. Предоставленная таким образом обратная связь является неинформативной в отношении нового задания: успех или неудача в тесте на аналитические способности не связаны с тем, насколько хорошо индивид знает страны и их столицы.

Установлено, что участники реагируют на негативную обратную связь, несмотря на то что задания не связаны друг с другом: группа участников, получившая негативный сигнал, значительно ниже оценила вероятность успеха на следующий день. Группа, получившая позитивный сигнал, не отличалась от контрольной. Все участники эксперимента учатся на разных факультетах МГУ им. М.В. Ломоносова. Характер восприятия ими обратной связи, возможно, отчасти обусловлен особенностями методики преподавания, принятыми во всем университете, такими как степень жесткости дедлайнов и критерии проверки письменных работ, специфика выбора образовательных траекторий и разделение на группы, степень открытости информации о результатах других студентов. Полученные выводы потенциально могут распространяться на другие высшие учебные заведения России с похожей системой, однако для этого требуется повторное проведение эксперимента на другой выборке.

Уверенность в собственном успехе положительно коррелирует с количеством ответов в тесте на соотнесение стран с их столицами: участники, более высоко оценившие вероятность своего успеха во второй день эксперимента, в среднем соотносят большее количество стран с их столицами. Предположительно уверенность в успехе влияет на выбираемый участниками уро-

вень усилий: участники, выше оценивающие свои шансы на победу, стараются вспомнить как можно больше правильных ответов. Участники, низко оценивающие вероятность оказаться среди победителей, не проходят тест до конца, оставляя часть полей пустыми.

Предложенный в работе дизайн эксперимента исключает возможность возникновения «эффекта горячей руки» за счет проведения эксперимента в два дня. Можно утверждать, что реакция индивидов на обратную связь, получаемую в условиях не связанных друг с другом заданий, не исчезает со временем — по крайней мере в течение двух дней, что важно учитывать при воздействии на самооценку и мотивацию студентов.

Описываемое воздействие неинформативной в отношении нового задания обратной связи на уверенность индивидов может привести к снижению результативности обучающихся, недооценивающих собственные способности. Необходимо принимать во внимание, что студенты реагируют в том числе на обратную связь, которая не несет дополнительной информации относительно новых испытаний. Следует искать более эффективные способы сообщать о результатах, не оказывая при этом неоправданного негативного влияния на самооценку и производительность отдельных групп.

7. Ограничения

Ограничения исследования связаны, во-первых, с недостаточной численностью участников. Большее количество наблюдений позволило бы проверить гипотезу о гетерогенности эффекта для групп участников, сформированных по разным признакам.

Во-вторых, в рамках лабораторного исследования сложно изучать влияние обратной связи на уровень прилагаемых участниками усилий: большая часть студентов обладает либо недостаточной, либо чрезмерной мотивацией. Так, участники, не заинтересованные в выигрыше, сразу пропускали задание с денежным вознаграждением и переходили к завершающему тесту. Другие, наоборот, старались пройти все раунды, не желая упустить потенциальный выигрыш. Чтобы усилия индивидов были более вариативны и зависели от их уверенности в себе, следует провести полевой эксперимент.

В-третьих, часть участников эксперимента обучаются вместе и осведомлены об академических успехах друг друга. Если участник из контрольной группы получит информацию о результатах теста на аналитические способности своего однокурсника, попавшего в группу воздействия, он может оценить свой результат, основываясь на своих представлениях о способностях однокурсника. В таком случае оказываемое воздействие (обратная связь) опосредованно влияет на уверенность участников контрольной груп-

пы (*spillover effect*). Мы просим каждого участника в письменном виде подтвердить согласие не разглашать результаты эксперимента до окончания исследования. Кроме того, участники сами выбирали удобный день для прохождения эксперимента, поэтому среди участников, выполнявших задания в одно время, находились студенты разных факультетов и разных курсов, не знакомые друг с другом, что существенно снижает вероятность получения информации об академических успехах других участников. При этом возможность распространения информации нельзя исключить полностью. Это не должно отразиться на основном выводе исследования: негативная обратная связь, получаемая в условиях не связанных друг с другом заданий, оказывает значимое воздействие на уверенность индивидов в собственных способностях. Если бы эффект отсутствовал, то полученный сигнал не оказывал бы влияния ни на группу контроля, ни на группу воздействия. Однако полученные нами оценки могут быть заниженными.

Приложения Приложение 1

Таблица П1. **Описательная статистика**

	Минимум	Максимум	Стандартное отклонение	Среднее значение	Медиана
Уверенность	0,00	1,00	0,30	0,69	0,75
Верно решенные матрицы Равена	2,00	16,00	3,09	8,79	9,00
Уверенность во второй день	0,00	1,00	0,32	0,50	0,50
Правильно переписанные строки	0,00	47,00	17,95	22,51	20,00
Правильно соотносенные столицы	0,00	12,00	2,55	5,86	6,00
Согласие с утверждением 1 (%)	0,00	100,00	31,34	48,03	50,00
Согласие с утверждением 2 (%)	0,00	100,00	28,76	47,47	45,00

Таблица П2. **Корреляционная матрица**

	Уверенность в первый день	Уверенность во второй день	Верно перепечатанные строки	Матрицы Равена	Столицы	Пол (М)
Уверенность в первый день		0,534***	0,086	0,068	0,188	0,030
Уверенность во второй день	0,534***		0,018	0,197	0,321**	-0,174
Верно перепечатанные строки	0,086	0,018		0,216*	0,314**	0,018

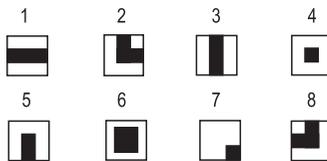
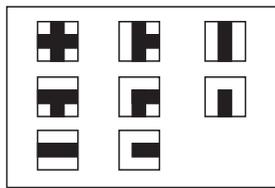
	Уверенность в первый день	Уверенность во второй день	Верно пере- печатанные строки	Матрицы Равена	Столицы	Пол (М)
Матрицы Равена	0,068	0,197	0,216*		0,190	-0,145
Столицы	0,188	0,321**	0,314**	0,190		-0,260*
Пол (М)	0,030	-0,174	0,018	-0,145	-0,260*	

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Приложение 2

Рис. П1. Пример матрицы Равена

Time left to complete this page: 0:27



Укажите номер недостающего элемента

Рис. П2. Пример одного из 50 раундов задания с денежным вознаграждением с активной кнопкой перехода к следующему заданию

Перепечатайте строчку 1, пропуская букву h, если она там есть

Вы можете пропустить последующие задания и приступить к прохождению теста

Предупреждение

Вы уверены в своем ответе? Перейти к следующему заданию?

Рис. П3. Первые четыре вопроса в задании на соотнесение стран и столиц

Time left to complete this page: 0:58

Выберите столицу для страны: Македония

Выберите столицу для страны: Черногория

Выберите столицу для страны: Гамбия

Приложение 3

Алгоритм генерации новых наблюдений с использованием ошибок, полученных с помощью метода бутстрэпа (*residual bootstrap*):

1. Оценивается модель регрессии на исходной выборке.
2. Полученные остатки модели

$$r_i = \log(\text{уверенность во второй день})_i - \widehat{\log(\text{уверенность во второй день})}_i$$

преобразовываются следующим образом:

$$r_i^* = \frac{\log(\text{уверенность во второй день})_i - \widehat{\log(\text{уверенность во второй день})}_i}{(1-h_i)^{1/2}}$$

где h_i — диагональный элемент матрицы оператора проектирования, отвечающий за вклад каждого наблюдения в прогноз. Скорректированные остатки $r_i^{new} = r_i^* - \bar{r}_i^*$ имеют равное нулю математическое ожидание и постоянную дисперсию — параметры соответствуют предпосылке о распределении случайных ошибок.

3. Скорректированные остатки случайным образом перераспределяются между наблюдениями. В работе перераспределение проводится 999 раз. Таким образом получены 999 квазивыборок с исходными регрессорами и новыми прогнозными значениями.
4. На каждой из квазивыборок регрессионная модель оценивается заново.

Стьюдентизированный доверительный интервал:

$$[\hat{\theta} - s.e.(\hat{\theta}) \cdot q_{0,975}; \hat{\theta} - s.e.(\hat{\theta}) \cdot q_{0,025}],$$

где $\hat{\theta}$ — оценка коэффициента для исходной выборки; $(\hat{\theta})$ — стандартная ошибка оценки коэффициента для исходной выборки; q — квантили соответствующего порядка бутстрэпируемой статистики $\frac{\hat{\theta} - \hat{\theta}_i^*}{s.e.(\hat{\theta}^*)}$, где $\hat{\theta}_i^*$ — оценка коэффициента для i -й квазивыборки.

Приложение 4

Таблица ПЗ. Анализ устойчивости результатов

Критерий создания переменной	Произведение переменных	Доля наблюдений, для которых произведение принимает значение, равное единице	<i>p-value</i> для теста на значимость коэффициента
1, если согласен с каждым из утверждений не менее чем на 20%	Согласие*Позитивный сигнал	0,16	0,54
	Согласие*Негативный сигнал	0,15	0,51
1, если согласен с каждым из утверждений не менее чем на 30%	Согласие*Позитивный сигнал	0,10	0,71
	Согласие*Негативный сигнал	0,11	0,74
1, если согласен с каждым из утверждений не менее чем на 50%	Согласие*Позитивный сигнал	0,03	0,68
	Согласие*Негативный сигнал	0,05	0,73

References

- Araujo de P., Lagos S. (2013) Self-Esteem, Education, and Wages Revisited. *Journal of Economic Psychology*, vol. 34, February, pp. 120–132. <https://doi.org/10.1016/j.joep.2012.12.001>
- Azmat G., Bagues M., Cabrales A., Iriberry N. (2019) What You Don't Know... Can't Hurt You? A Natural Field Experiment on Relative Performance Feedback in Higher Education. *Management Science*, vol. 65, no 8, pp. 3714–3736. <https://doi.org/10.1287/mnsc.2018.3131>
- Azmat G., Iriberry N. (2010) The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students. *Journal of Public Economics*, vol. 94, no 7–8, pp. 435–452. <https://doi.org/10.1016/j.jpubeco.2010.04.001>
- Banerjee R., Gupta N.D., Villeval M.C. (2020) Feedback Spillovers across Tasks, Self-Confidence and Competitiveness. *Games and Economic Behavior*, vol. 123, September, pp. 127–170. <https://doi.org/10.1016/j.geb.2020.07.002>
- Barankay I. (2012) Rank Incentives: Evidence from a Randomized Workplace Experiment. *Business Economics and Public Policy Papers*. Available at: <https://api.semanticscholar.org/CorpusID:5765998> (accessed 20 July 2024).
- Barron K. (2021) Belief Updating: Does the 'Good-News, Bad-News' Asymmetry Extend to Purely Financial Domains? *Experimental Economics*, vol. 24, no 1, pp. 31–58. <https://doi.org/10.1007/s10683-020-09653-z>
- Bénabou R., Tirole J. (2016) Mindful Economics: The Production, Consumption, and Value of Beliefs. *Journal of Economic Perspectives*, vol. 30, no 3, pp. 141–164. <https://doi.org/10.1257/jep.30.3.141>
- Brandts J., Groenert V., Rott C. (2015) The Impact of Advice on Women's and Men's Selection into Competition. *Management Science*, vol. 61, no 5, pp. 1018–1035. <https://doi.org/10.1287/mnsc.2013.1877>
- Burks S.V., Carpenter J.P., Goette L., Rustichini A. (2013) Overconfidence and Social Signaling. *Review of Economic Studies*, vol. 80, no 3, pp. 949–983. <https://doi.org/10.1093/restud/rds046>
- Buser T. (2016) The Impact of Losing in a Competition on the Willingness to Seek Further Challenges. *Management Science*, vol. 62, no 12, pp. 3439–3449. <https://doi.org/10.1287/mnsc.2015.2321>

- Buser T., Gerhards L., van der Weele J. (2018) Responsiveness to Feedback as a Personal Trait. *Journal of Risk and Uncertainty*, vol. 56, no 2, pp. 165–192. <https://doi.org/10.1007/S11166-018-9277-3/TABLES/7>
- Buser T., Yuan H. (2016) *Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad*. Tinbergen Institute Discussion Paper no 16-096/I. <https://doi.org/10.2139/ssrn.2867346>
- Carrillo J.D., Mariotti T. (2000) Strategic Ignorance as a Self-Disciplining Device. *Review of Economic Studies*, vol. 67, no 3, pp. 529–544. <https://doi.org/10.1111/1467-937x.00142>
- Davison A.C., Hinkley D.V. (1997) *Bootstrap Methods and their Application in Bootstrap Methods and their Application*. Cambridge: Cambridge University. <https://doi.org/10.1017/cbo9780511802843>
- Drouvelis M., Paiardini P. (2022) Feedback Quality and Performance in Organizations. *Leadership Quarterly*, vol. 33, no 6, Article no 101534. <https://doi.org/10.1016/j.leaqua.2021.101534>
- Eil D., Rao J.M. (2011) The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics*, vol. 3, no 2, pp. 114–138. <https://doi.org/10.1257/mic.3.2.114>
- Ertac S. (2011) Does Self-Relevance Affect Information Processing? Experimental Evidence on the Response to Performance and Non-Performance Feedback. *Journal of Economic Behavior and Organization*, vol. 80, no 3, pp. 532–545. <https://doi.org/10.1016/j.jebo.2011.05.012>
- Filippin A., Gioia F. (2018) Competition and Subsequent Risk-Taking Behaviour: Heterogeneity across Gender and Outcomes. *Journal of Behavioral and Experimental Economics*, vol. 75, August, pp. 84–94. <https://doi.org/10.1016/j.socec.2018.05.003>
- Gill D., Prowse V. (2012) A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *American Economic Review*, vol. 102, iss. 1, pp. 469–503. <https://doi.org/10.1257/aer.102.1.469>
- Gilovich T., Vallone R., Tversky A. (1985) The Hot Hand in Basketball: On the Misperception of Random Sequences. *Cognitive Psychology*, vol. 17, no 3, pp. 295–314. [https://doi.org/10.1016/0010-0285\(85\)90010-6](https://doi.org/10.1016/0010-0285(85)90010-6)
- Grossman Z., Owens D. (2012) An Unlucky Feeling: Overconfidence and Noisy Feedback. *Journal of Economic Behavior and Organization*, vol. 84, no 2, pp. 510–524. <https://doi.org/10.1016/j.jebo.2012.08.006>
- Hermes H., Huschens M., Rothlauf F., Schunk D. (2019) *Motivating Low-Achievers – Relative Performance Feedback in Primary Schools*. NHH Dept. of Economics Discussion Paper no 14. <https://doi.org/10.2139/ssrn.3411429>
- Huang L., Murad Z. (2019) *Feedback Spillover Effect on Competitiveness Across Unrelated Tasks*. Available at: <https://ssrn.com/abstract=3436581> or <http://dx.doi.org/10.2139/ssrn.3436581> (accessed 12 August 2024).
- Karlsson N., Loewenstein G., Seppi D. (2009) The Ostrich Effect: Selective Attention to Information. *Journal of Risk and Uncertainty*, vol. 38, no 2, pp. 95–115. <https://doi.org/10.1007/sl>
- Kuhnen C. M., Tymula A. (2012) Feedback, Self-Esteem, and Performance in Organizations. *Management Science*, vol. 58, no 1, pp. 94–113. <https://doi.org/10.1287/mnsc.1110.1379>
- Linhares L.V., Torres A.R.R., Pereira C.R. (2022) Live by the Sword, Die by the Sword: Measuring Belief in a Just World with Popular Sayings. *Personality and Individual Differences*, vol. 195, September, Article no 111673. <https://doi.org/10.1016/j.paid.2022.111673>
- Lovász A., Cukrowska-Torzewska E., Rigó M., Szabó-Morvai Á. (2022) Gender Differences in the Effect of Subjective Feedback in an Online Game. *Journal of Behavioral and Experimental Economics*, vol. 98, June, Article no 101854. <https://doi.org/10.1016/j.socec.2022.101854>

- Mammen E. (1992) Bootstrap, Wild Bootstrap, and Asymptotic Normality. *Probability Theory and Related Fields*, vol. 93, no 4, pp. 439–455. <https://doi.org/10.1007/BF01192716>
- Mayraz G. (2011) *Wishful Thinking*. Available at: <https://ssrn.com/abstract=1955644> or <http://dx.doi.org/10.2139/ssrn.1955644> (accessed 20 July 2024).
- McLeod C., Ponesse J.E. (2008) Infertility and Moral Luck: The Politics of Women Blaming Themselves for Infertility IJFAB. *International Journal of Feminist Approaches to Bioethics*, vol. 1, no 1, pp. 125–144. <https://doi.org/10.1353/IJF.0.0001>
- Mezulis H.A., Abramson Y.L., Hyde S.J., Hankin L.B. (2004) Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias. *Psychological Bulletin*, vol. 130, no 5, pp. 711–747. <http://dx.doi.org/10.1037/0033-2909.130.5.711>
- Miller D.T., Ross M. (1975) Self-Serving Biases in the Attribution of Causality: Fact or Fiction? *Psychological Bulletin*, vol. 82, no 2, pp. 213–225. <https://doi.org/10.1037/h0076486>
- Möbius M.M., Niederle M., Niehaus P., Rosenblat T.S. (2022) Managing Self-Confidence: Theory and Experimental Evidence. *Management Science*, vol. 68, no 11, pp. 7793–8514. <https://doi.org/10.1287/mnsc.2021.4294>
- Moore D.A., Healy P.J. (2008) The Trouble with Overconfidence. *Psychological Review*, vol. 115, no 2, pp. 502–517. <https://doi.org/10.1037/0033-295X.115.2.502>
- Raven J., Raven J.C., Court J.H. (1998) *Raven Manual: Section 4, Advanced Progressive Matrices*. Oxford: Oxford Psychologists.
- Schwardmann P., van der Weele J. (2019) Deception and Self-Deception. *Nature Human Behaviour*, vol. 3, no 10, pp. 1055–1061. <https://doi.org/10.1038/s41562-019-0666-7>
- Shah P., Harris A.J.L., Bird G., Catmur C., Hahn U. (2016) A Pessimistic View of Optimistic Belief Updating. *Cognitive Psychology*, vol. 90, November, pp. 71–127. <https://doi.org/10.1016/j.cogpsych.2016.05.004>
- Sharot T. (2011) *The Optimism Bias: A Tour of the Irrationally Positive Brain Choice*. New York, NY: Pantheon Books.
- Shastry G.K., Shurchkov O., Xia L.L. (2020) Luck or Skill: How Women and Men React to Noisy Feedback. *Journal of Behavioral and Experimental Economics*, vol. 88, October, Article no 101592. <https://doi.org/10.1016/j.socecon.2020.101592>
- Steeh A., Höffler T., Höft L., Parchmann I. (2021) First Steps toward Gender Equity in the Chemistry Olympiad: Understanding the Role of Implicit Gender-Science Stereotypes. *Journal of Research in Science Teaching*, vol. 58, no 1, pp. 40–68. <https://doi.org/10.1002/tea.21645>
- Svenson O. (1981) Are We All Less Risky and More Skillful Than Our Fellow Drivers? *Acta Psychologica*, vol. 47, no 2, pp. 143–148. [https://doi.org/10.1016/0001-6918\(81\)90005-6](https://doi.org/10.1016/0001-6918(81)90005-6)
- Sweeny K., Melnyk D., Miller W., Shepperd J.A. (2010) Information Avoidance: Who, What, When, and Why. *Review of General Psychology*, vol. 14, no 4, pp. 340–353. <https://doi.org/10.1037/a0021288>
- Tran A., Zeckhauser R. (2012) Rank as an Inherent Incentive: Evidence from a Field Experiment. *Journal of Public Economics*, vol. 96, no 9–10, pp. 645–650. <https://doi.org/10.1016/j.jpubeco.2012.05.004>
- Verna M. (2002) Understanding Gender Inequity in America: Interviews with Academic Olympians. *Research in Education*. Available at: https://www.academia.edu/27019003/Understanding_gender_inequity_in_America_Interviews_with_academic_Olympians (accessed 20 July 2024).
- Villeval M.C. (2020) Performance Feedback and Peer Effects: A Review. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3550667>
- Wilcox R.R., Keselman H.J. (2003) Modern Robust Data Analysis Methods: Measures of Central Tendency. *Psychological Methods*, vol. 8, no 3, pp. 254–274. <https://doi.org/10.1037/1082-989X.8.3.254>

- Wozniak D., Harbaugh W.T., Mayr U. (2016) *The Effect of Feedback on Gender Differences in Competitive Choices*. Available at SSRN: <https://ssrn.com/abstract=1976073> or <http://dx.doi.org/10.2139/ssrn.1976073> (accessed 12 August 2024).
- Wozniak D., Harbaugh W. T., Mayr U. (2014) The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices. *Journal of Labor Economics*, vol. 32, no 1, pp. 161–198. <https://doi.org/10.1086/673324>
- Wu C.F.J. (2007) Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *The Annals of Statistics*, vol. 14, no 4, pp. 1261–1295. <https://doi.org/10.1214/aos/1176350142>
- Zimmermann F. (2020) The Dynamics of Motivated Beliefs. *American Economic Review*, vol. 110, no 2, pp. 337–361. <https://doi.org/10.1257/aer.20180728>