ISSN (Print) 0473-453X
Discussion Paper No. 1242
ISSN (Online) 2435-0982

When Learning Together Goes Wrong: Negative Peer Effects in Online Learning

Shohei Yamamoto
Shuma Iwatani

Revised July 2024
April 2024

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

When Learning Together Goes Wrong: Negative Peer Effects in Online Learning

Shohei Yamamoto[1] and Shuma Iwatani[2]

Abstract

This research examined the impacts of peer perseverance on learning outcomes and investigated how this influence varies with the partners' skill levels, either higher or similar to the subjects, across two experiments resembling online learning platforms. Study 1 recruited active English learners, while Study 2 involved subjects who had not engaged in English studies for more than six months. The results of both experiments revealed negative rather than positive peer effects. The subjects ceased studying earlier and displayed lower performance when learning with less perseverant peers, compared to when studying alone. This pattern was observed when paired with similarly skilled peers in Study 1 and with higher-skilled peers in Study 2. Further analysis indicated that the negative peer effects predominantly originated from subjects with lower levels of motivation. Additionally, it was shown that social proximity could foster positive effects when peers possess similar skills and higher perseverance levels.

Keywords: peer effects, perseverance, performance, online learning

[1] Graduate School of International Corporate Strategy, Hitotsubashi University, Tokyo, Japan (ORCID: 0000-0002-8136-6842). Correspondence concerning this article should be addressed to Shohei Yamamoto. Email: syamamoto@ics.hub.hit-u.ac.jp

[2] School of Sociology, Kwansei Gakuin University, Hyogo, Japan (ORCID: 0000-0002-1836-0456). Email: iwatani.shuma@kwansei.ac.jp

When Learning Together Goes Wrong: Negative Peer Effects in Online Learning

The advent of the digital age has revolutionized various sectors, with the education industry being a prominent beneficiary. The emergence and subsequent ubiquity of online learning platforms have provided an opportunity to transcend geographical boundaries and time constraints, offering a new dimension to the traditional learning methods. These platforms cater a diverse range of learning needs, from programming (e.g., Codecademy) and language acquisition (e.g., Duolingo) to professional certifications courses (e.g., Udemy). The online learning market has been expanding at an unprecedented rate. As of 2022, the industry was valued at $198.2 billion and is projected to reach a value of $602.0 billion by 2030 (Vantage Market Research, 2023). Duolingo, a leading language learning platform, reported a 23% increase in monthly active users, rising to 49.2 million in 2022.

Despite the surging popularity and the seemingly limitless opportunities offered by online learning platforms such as Coursera, Skillshare, LinkedIn Learning, as well as massive open online courses, a recurring issue that plagues the sector is the high attrition rate (Eriksson et al., 2017; Narayanasamy & Elçi, 2020). Indeed, research suggests that student attrition rates on online courses are generally higher than those in traditional classroom-based courses (Levy, 2007; Tello, 2007).

High dropout rates would undermine the potential of these platforms and pose a significant challenge to realizing the full potential of online learning. To tackle the problem, some platforms have introduced social features for learners to engage more. As an example, Duolingo has implemented several such features including leaderboards and a function to follow other users.[1] One particularly intriguing feature is “Friends Quest,” where users are randomly paired up with one of

[1] The following function allows users to follow other users. Duolingo claims that learners who follow other users on the platform are 5.6 times more likely to finish their course (Zabell, 2023). Leaderboards list the top engaging users and enable learners to monitor their weekly progress compared to other learners worldwide. Users can ascend the leaderboard through active engagement with the platform and gauge their standings relative to the users they follow.

their Duolingo friends each week. They are assigned a challenge such as completing a certain number of lessons to earn rewards.

As illustrated in the examples above, online learning platforms have been striving to strengthen the engagement of users by promoting interactions among users. This raises an essential question regarding the impact of such interactions on individual perseverance and performance. The following section will provide a comprehensive literature review to explore this issue.

Effects of Peer's Perseverance

Literature suggests that individuals are influenced by the perseverance of their peers, with positive effects stemming from high perseverance and negative effects from low perseverance (Battaglini et al., 2017; Buechel et al., 2014; Gerhards & Gravert, 2020; Golsteyn et al., 2021). Battaglini et al. (2005) developed an economic model to explain such peer effects. The model posits that witnessing peers successfully overcome self-control challenges boosts individuals' confidence and belief in their own ability to manage temptations. Conversely, observing peer failures undermines individuals' belief in their potential for success.

Peer Effects from Similar Peers

The extent of peer effects varies based on peer characteristics, with individuals more likely to be influenced by peers they resemble (Battaglini et al., 2005; Booij et al., 2017; Carrell et al., 2013; Hanna et al., 2013). Social learning theory supports this concept, showing that individuals are more inclined to imitate the behaviors of peers who are similar (e.g., skills, status, or age) (Hanna et al., 2013). Furthermore, Battaglini et al. (2005) posited in their model that individuals acquire insights about their own perseverance level from similar others.

The stronger influence from similar peers may be attributed to perceived social closeness. Rosaz et al. (2016) suggested that similar quitting times for tasks between partners were due to reduced social distances among them. Bicchieri et al. (2022) examined the effects of peer influence on

socially desirable and undesirable behaviors, analyzing the role of social proximity. Their findings indicate that in the absence of social proximity, peers predominantly influenced undesirable behaviors but not desirable behaviors. However, when social proximity was present, both desirable and undesirable behaviors were significantly impacted.

This concept of social proximity may also elucidate the findings of Bandiera et al. (2010), who found that worker productivity was influenced when friends, rather than mere acquaintances, were working nearby. Additionally, Lomi et al. (2011) demonstrated that students with similar skill levels are more likely to form friendships and advisory relationships, further reinforcing the link between similarity and enhanced peer influence.

Peer Effects from High-skilled Peers

Peers with higher skill levels also exert substantial influence. According to social learning theory, individuals are more inclined to imitate behaviors of peers who are perceived as attractive, successful, or of high status (Hanna et al., 2013). Individuals are more likely to pay attention to and learn from those skilled at a task (Bandura, 1986) and copy successful strategies (Mesoudi, 2008; Mesoudi & O'Brien, 2008). Similarly, Henrich & Gil-White (2001) identified a *prestigious bias*, where individuals selectively copy prestigious peers. This phenomenon of copying successful, prestigious, or high-skilled people is known as *indirect bias*. Adopting the traits of such individuals is seen as increasing one's own chances of success (Boyd & Richerson, 1985; Mesoudi & O'Brien, 2008). Research beyond social comparison theory also showed significant influences from high-skilled peers, where being observed by them can increase productivity due to self-image concerns (Boyd & Richerson, 1985; Bursztyn & Jensen, 2017; Mas & Moretti, 2009), while some studies report negative effects from high-skilled peers among low ability individuals (e.g., Feld & Zölitz, 2017).

Field Studies and Laboratory Experiments

Our literature review indicates that the perseverance of learning partners with similar or higher skill levels than the individuals significantly impacts the magnitude of peer effects. However, these studies are likely subject to biased statistical estimates of peer effects. Feld and Zölitz (2017) noted that estimates of peer effects vary widely across research possibly because field studies on peer effects commonly face identification problems (Angrist, 2014; Manski, 1993; Sacerdote, 2014).[2] Several studies have rigorously addressed these issues by exogenously assigning individuals to peer groups and using pre-treatment characteristics to measure peer ability (Carrell et al., 2009, 2013; Duflo et al., 2011; Lyle, 2007). However, despite these sophisticated approaches, the estimates of peer effects can still be prone to biases (Angrist, 2014; Feld & Zölitz, 2017).

Conversely, experimental studies in the lab can circumvent this issue by having subjects observe peers before engaging in separate real-effort tasks that hold little to no intrinsic value (e.g., Beugnot et al., 2019; Buechel et al., 2014, 2018; Georganas et al., 2015; Gerhards & Gravert, 2020) However, these real-effort tasks lack realism and generalizability to real-world settings.

Research Questions

Drawing on the previous literature mentioned above, we anticipate that individuals are affected by the perseverance of their peers. The impact is expected to be strong from high-skilled peers or peers with similar skills to individuals. However, the differential impact of these skill levels remains unclear. Therefore, this study seeks to address the following research questions:

- How does the perseverance of peers affect both an individual's own perseverance and performance?

[2] Two major obstacles in identifying peer influences are the selection and reflection issues. The selection issue arises because peer groups often form naturally, making it challenging to separate the influence of peers from the effects of group selection. The reflection issue occurs because it is difficult to differentiate the impact that peers have on an individual from the influence that the individual has on their peers when these effects occur simultaneously.

- Does this effect vary when peers possess higher skills than the individual compared to when peers have similar skills?

We used a novel experimental method designed to mirror realistic online learning environments with intrinsic value, free from the identification problems typically encountered in field studies. This approach enables us to identify and estimate unbiased peer effects while ensuring strong external validity.

Overview of Two Experiments

We recruited Japanese residents and conducted experiments,[3] offering subjects opportunities to learn English vocabulary online which resemble a typical English online learning environment. The choice of English vocabulary learning as the context was driven by the expectation that subjects would have some intrinsic value in learning English.

In the experiment, we compared the amount of effort exerted in studying English as well as test scores between the subjects studying alone (single condition) and those studying with a partner who was pre-recruited in advance of the main study (pair condition). In the pair condition, individuals were informed about their partners' level of perseverance, specifically whether the partner continued learning or not.

This approach enabled us to clearly assess the causal influence of the partners' perseverance on objective measures of subjects' perseverance as well as their performance, unlike typical studies that only measure perseverance or rely on self-reported measures (e.g., Buechel et al., 2014).

[3] Japan serves as a suitable context for this research due to its large market for English language education. The immense interest in English proficiency within the country is evident in the participation rates for the Eiken (an English proficiency exam inaugurated in 1963) and Test of English for International Communication (TOEIC). In 2022 alone, over 4 million individuals took the Eiken test, while more than 2 million participated in the TOEIC exam (including the "TOEIC Listening & Reading Test" and "TOEIC Speaking & Writing Tests"). These figures underscore the significant demand for English language proficiency in Japan.

We conducted two studies. In Study 1, we recruited subjects who were currently studying English. In Study 2, we recruited subjects who had not engaged in studying for more than six months. The subjects in Study 2 were less skilled and did not enjoy studying compared to the subjects in Study 1.

We hypothesize that, in such scenarios, the duration of time committed to studying is likely to be more closely aligned with that of their peers. Specifically, individuals' perseverance is expected to increase when paired with high-perseverance peers and decrease when paired with low-perseverance peers, relative to studying alone. The distinction in the strength of influence between peers with similar versus higher skills presents an intriguing empirical question due to the lack of clear predictions in existing literature. Additionally, we also anticipate that this influence intensifies if individuals perceive a sense of social proximity to their peers.

In both experiments, we observed negative peer effects stemming from partners with lower levels of perseverance. Both the number of words studied and the test scores were lower when partners possessed low level of perseverance, compared to outcomes when studying independently. The magnitude of the effect was substantial. The number or words studied dropping by about 11 to 15 words in both studies, where they could study up to 95 words. In terms of test scores (maximum of 30 points), there was a decrease of approximately 1.8 points in Study 1. These findings were consistent with prior research indicating negative peer influences across various contexts, including cooperation, dishonesty, and prosocial behaviors (Bicchieri et al., 2022; Isler & Gächter, 2022; Rauhut, 2013; Thöni & Gächter, 2015). In contrast, positive peer effects were not observed on average. This asymmetry in peer effects is also consistent with previous studies (Isler & Gächter, 2022; Thöni & Gächter, 2015).

In addition, the source of negative effects varied between the two studies. In Study 1, negative effects from partners with similar skill levels were observed. In contrast, in Study 2, negative effects

emerged among subjects paired with high-skilled partners, who were less proficient than those in Study 1. The interpretation of the findings will be discussed in the General Discussion and Conclusion section.

In addition to this, we found that perceived social proximity and motivation toward learning can affect the strength of peer effects. Our analysis shows that social proximity, as measured by the Inclusion of Other in Self (IOS) scale (Aron et al., 1992), may enhance positive effects among partners with similar skill levels and higher perseverance. This result is consistent with prior findings indicating stronger peer effects among individuals with closer social connections (Bandiera et al., 2010).

Moreover, when segmenting the sample based on the level of intrinsic motivation for learning the English language, the findings demonstrated that less intrinsically motivated subjects experienced substantial negative peer effects from low perseverance partners, while more motivated subjects did not. This observation aligns with the self-determination theory, which posits that intrinsically motivated individuals rely less on external incentives or pressure, deriving satisfaction and joy from their own endeavors (Ryan & Deci, 2020).

Study 1: Influence of Partners' Perseverance on English Vocabulary Learning

Study 1 was designed to examine the clear causal influences of learning partners' perseverance on individuals' own perseverance and performance, and how this influence varies with the partners' skill levels, either being higher or similar to those of the subjects.

Before the main study, we separately conducted a study to recruit potential partners who were matched with the subjects in the main study. We recruited four potential partners, each possessing different characteristic factors: two distinguished by gender (male and female) and two

distinguished by their level of English skills (high and middle).[4] We initially asked them to study six and then 95 English words. Additionally, we instructed them to send emojis from time to time to other subjects to motivate their learning. We used the data of high-skilled female and middle-skilled female as potential partners in the main studies.[5]

Method

Subjects. In Study 1, we recruited 400 subjects through Lancers, Inc. (https://www.lancers.jp). The criteria for our subjects were as follows; (1) Japanese residents, (2) currently studying English and (3) having English skills less than the C1 level (less than a TOEIC score of 950). We excluded 26 subjects whose vocabulary level exceeded the scope of this experiment as described in the following section. Consequently, we obtained a demographically diverse sample of 373 subjects in total (48% female, M_{age} = 38.8 years, age range: 18-72 years). Of these subjects, 52% are fully employed, 14% hold part-time positions, 5% are students, and the remaining 29% fall into other categories. The subjects received a fixed fee of 600 yen. The duration of the experiment varied depending on how many English words they chose to learn, with the median duration of approximately 48 minutes.

Experimental Design. Subjects were randomly assigned to one of the five conditions: single or one of four different pair conditions. The subjects in the single condition did not have any learning partners, while the subjects in the pair conditions had learning partners (whom we had recruited before conducting this experiment as mentioned above). Furthermore, the characteristics of partners were manipulated in the pair condition: English skills (high/similar to the subject), and their perseverance level regarding study time (quitter/finisher).

[4] Test of English for International Communication (TOEIC) score of 955 is considered indicative of as high English skills and middle TOEIC score of 615 is considered indicative of middle English skills.
[5] We used female as potential partners in this experiment. This is because subjects felt their peers more beneficial in learning English when paired with female partners, compared to male partners, according to the questionnaire in the pretest.

Procedure. Subjects took part in the five parts of the experiment: pre-learning part, pre-study questionnaire and pairing announcement part, mandatory learning part, voluntary learning part and test part (see Table 1 for a summary of the procedure).

Pre-learning Part. In the pre-learning part, subjects were presented with English words and simply answered whether they knew the words. They chose the option between "I know the word" and "I did not know the word at all" (they were instructed to choose "I know the word" even if they had some familiarity with the word). We selected 500 English words for the subjects from Shimazu's vocabulary book (Shimazu, 2021). The level of the selected words ranged from B1 to C2 levels in Common European Framework of Reference for Languages. We excluded 26 subjects who claimed to know more than 400 words in this part, as their English proficiency was estimated to be higher than the C1 level.

Pre-Study Questionnaire and Pairing Announcement Part. After the number of words that they did not know reached 100, they proceeded to the pre-study questionnaire and pairing announcement part. Here, they completed a demographic survey that inquired about various attributes, including age, gender, education, marital status, and employment status. The survey also asked about their primary motivations for learning English as well as their self-assessed level of English proficiency, which ranged from beginner (1) to expert (5).[6] Additionally, they were instructed to choose their own avatar like the ones in typical smartphone apps.

[6] The options for the reasons for learning English include: for work, for travel, to communicate with international people, to help foreigners in need, to gain common sense, because it is cool, to enjoy foreign movies and other media, for the sake of children, and other reasons. The subjects who selected 5 on the English proficiency question were not allowed to continue the experiment as per the instructions.

Table 1

Summary Procedure of the Experiment

<table>
<tr><th></th><th>Single condition</th><th colspan="4">Pair condition</th></tr>
<tr><th rowspan="2">Part</th><th rowspan="2"></th><th colspan="2">High-skilled</th><th colspan="2">Similar-skilled</th></tr>
<tr><th>Finisher</th><th>Quitter</th><th>Finisher</th><th>Quitter</th></tr>
<tr><td>1. Pre-learning</td><td colspan="5">Answered if they recognized given English words
Answered demographics and other related questions</td></tr>
<tr><td>2. Pre-study questionnaire and pairing announcement</td><td>No partner information provided.</td><td colspan="4">Received information on partner’s English skills
Informed that partner answered 3 more words correctly.
Informed that partner answered the same number of words correctly.</td></tr>
<tr><td>3. Mandatory learning</td><td colspan="5">Studied 5 out of 100 unfamiliar English words from pre-learning.</td></tr>
<tr><td>4. Voluntary learning</td><td>Had the option to study English words alone. Could stop anytime after the first question.</td><td colspan="4">Had the option to study English words with the partner. Could stop anytime after the first question.
Informed that their partner continued to study all questions (up to 95).
Informed that their partner stopped after the 6th question.
Informed that their partner continued to study all questions (up to 95).
Informed that their partner stopped after the 6th question.</td></tr>
<tr><td>5. Test</td><td colspan="5">Took a vocabulary test on 30 words they identified as unfamiliar during pre-learning.</td></tr>
</table>

Immediately after the questions, subjects in the pair conditions were notified that they were paired with their learning partner. They were also informed of the partner's characteristics, including gender, a randomly generated name, and English proficiency skills (see Figure 1 for the screenshot). At this point, we implemented our first manipulation of the partners' skill. Subjects in the similar-skilled condition were informed that the partners answered the same number of words correctly as they did for the 10 questions in the pre-learning part. On the other hand, subjects in the high-skilled condition were informed that the partners answered 3 more words correctly than they did for the 10 questions. We deliberately picked 10 questions in the pre-learning part to avoid deception problems. For example, we selected 10 questions where the subjects in the high-skilled condition correctly answered 5 questions, and 10 questions where their partner correctly answered 8 questions. Subjects in the single condition did not receive any information about partners.

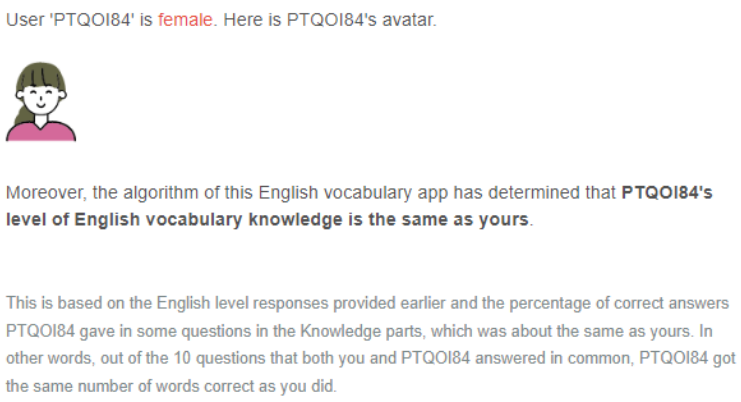

Figure 1. The Information of the Partner in the Instructions (Similar-skilled Partner Condition).

Mandatory Learning Part. Subsequently, they proceeded to the mandatory learning part. Following the methodology in Rosaz et al. (2016), this part was designed to let subjects to familiarize the procedure, as well as the level of difficulty and duration of our experimental task. Here, they must learn five words out of 100 words that they answered "I did not know the word at

all" in the pre-learning part. In this part, each question displayed an English word above five Japanese translation options (Figure 2).[7] Subjects were asked to select one of the options that correctly translated the given English word within 10 seconds. After that, the feedback was shown and they were asked to memorize it for five seconds. This process was repeated five times, resulting in the study of five words in this part. There was no break while doing this part. If they did not choose any option within the time period, they had a second chance to answer the same question within 10 seconds. Failure to choose an option during this additional opportunity resulted in the termination of their participation in the study.

Voluntary Learning Part. Upon completing the mandatory learning part, they went to the voluntary learning part. The procedure here was akin to the mandatory learning part, but with notable distinctions. Specifically, subjects had the option to cease their learning at any point after answering one question. They could study up to 95 additional words. To ensure comprehension of these conditions, subjects were required to answer a comprehension question, clarifying their understanding that they could discontinue the learning task at their discretion after answering the first question. After they correctly answered this question, they were allowed to start learning.

In this part, the type of information received differed depending on whether subjects were in the single or paired conditions. Those in the single condition continued their learning alone, as in the mandatory section. On the other hand, subjects in the paired condition engaged in learning with information of their partners. They received emojis from their partner (refer to Figure 2) and had the opportunity to send emojis to their partner. This procedure aimed to foster a sense of learning with their partners.

[7] Multiple-choice questions are a prevalent method for assessing vocabulary proficiency. Presenting answer choices in the first language, which is Japanese in our study, is efficient (Nation, 2001).

At this stage, we introduced the second manipulation concerning partners' perseverance level (quitter / finisher). Subjects in the quitter condition were notified that their partner had ceased studying after six questions. Conversely, those in the finisher condition were informed that the partner continued to study questions.

Test Part. After the voluntary learning part, subjects transitioned to the test part. This part consisted of a vocabulary test featuring 30 words that subjects had previously identified as "I did not know the word at all" in the pre-learning part.

Following the test, subjects in the paired conditions responded to survey questions concerning their perceptions of their partners. The survey questions encompassed several aspects, including: How was the amount of words learned by the partner (1: very little – 10: very much)?; How do you think your partner's level of English vocabulary compares to yours (1: much lower – 7: much higher)?; Do you think it is beneficial to have the partner on your learning (0: not beneficial at all – 10: very beneficial)? and Inclusion of Other in Self Scale (IOS; Aron et al., 1992) measuring how close the subjects felt with their partner (1: no overlap – 7: most overlap); How did the amount of words learned by the partner's influence the amount of words learned by you (1: decreased a lot – 7: increased a lot)? Lastly, we also asked how fun learning English words in this study was (with an 11-point scale).

Figure 2. The Screenshot of an Example Question in the Voluntary Learning Part (Pair Condition).

In the upper right corner, the learning status of the partner (PTQOI84) and their sent emoji are displayed. Subjects are required to select the correct Japanese translation of the presented English word from five given options.

Results

Manipulation Check

The manipulations of the partners' skills and perseverance in this study were successful. Subjects in the high-skilled condition believed their partner had significantly higher vocabulary levels (M = 5.39, SD = 1.24) compared to those in the similar-skilled condition (M = 4.25, SD = 1.02; $F_{(1,301)}$ = 76.50, $p < .001$, $\eta^2 = .20$). Subjects in the finisher conditions perceived that their partner studied more (M = 7.02, SD = 2.26) compared to those in the quitter conditions (M = 3.68, SD = 2.42; $F_{(1,301)} = 153.77$, $p < .001$, $\eta^2 = .34$).

Descriptive Analysis

Table 2 shows the selected characteristics of the subjects for each treatment. The variable *known_words* represents the number of words subjects knew in the pre-learning part, *eng_level* denotes their self-assessed level of English proficiency, and the dummy variable *female* indicates if

subjects are female. The variables between the treatments are not statistically significant indicating that the randomization works well (all p-values above 0.1).

Table 2

Descriptive Statistics (Study 1)

	Similar-quitter	**High-quitter**	**Similar-finisher**	**High-finisher**	**Single**	**Statistic test**	**p-value**
known_words	80.00	86.81	90.48	81.01	83.94	$F_{(4,368)} = 0.27$	$p = .90$
	(70.82)	(70.10)	(88.42)	(65.17)	(59.70)		
eng_level	1.80	1.87	1.89	2.00	1.89	$F_{(4,368)} = 0.48$	$p = .75$
	(0.84)	(0.88)	(0.97)	(0.89)	(0.84)		
age	38.64	37.96	38.84	40.51	38.27	$F_{(4,368)} = 0.61$	$p = .66$
	(10.30)	(11.46)	(11.47)	(10.06)	(10.50)		
female	0.45	0.52	0.51	0.45	0.47	$\chi_{(4)} = 1.31$	$p = .86$
N	74	79	79	71	70		

Notes: Standard errors in parentheses. The variable *known_words* represents the number of words subjects knew in the pre-learning part; *eng_level* denotes their self-assessed level of English proficiency; and *female* indicates the proportion of female subjects.

The top of Figure 3 illustrates that the number of studied words in the voluntarily part across different conditions. Subjects had their inherent motivation to engage with English vocabulary learning. They learned 38.31 words on average in the voluntary learning part, which was significantly larger than 1 (t (372) = 18.67, $p < .001$). In addition, the average enjoyment rating for this experiment was 7.52 (on a scale from 0 to 10), suggesting a high level of enjoyment.

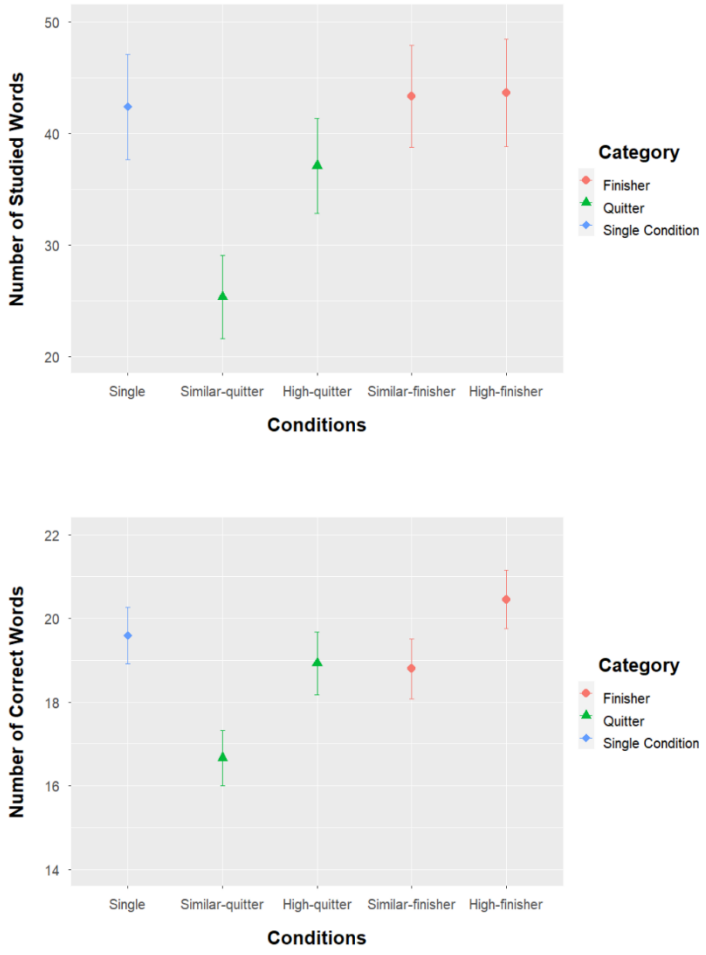

Figure 3. The Number of Studied Words in the Voluntarily Learning Part (Top) and the Number of Correct Words in the Test Part (Bottom) for Each Treatment (Study 1).

Means are represented by three distinct shapes: a circle for Finisher, a triangle for Quitter, and a diamond for Single condition. Standard error bars are included.

Regression Analysis

Perseverance. We conducted our analysis with three different models employing OLS regression. In the first model, we assessed the impact of learning partners' perseverance on the number of words studied, comparing partners with low and high perseverance against the single condition baseline (see column 1 at the top of Table 3). The dummy variable *quitters* is assigned a

value of 1 if the partner's perseverance is low and 0 otherwise, while *finishers* is set to 1 for high perseverance. In the second model, we conducted the analysis instead of using the dummy variables of *quitters* and *finishers*, we utilized four treatment dummy variables: *high-finisher*, *high-quitter*, *similar-finisher*, and *similar-quitter*, which reflect a combination of the partners' English skill levels (high or similar) and their perseverance (quitter or finisher) (see column 2). In the third model, we incorporated demographic variables into the second model. Demographic variables include age, gender, and college—defined by a dummy variable that takes the value of 1 if the subject has a college degree—as well as a variable *known_words* for the number of known words in the pre-learning part (see column 3).

Next, we repeated the analysis of these three models using Tobit regression to account for the constraint that subjects could not study fewer than 1 or more than 95 words and could not correctly answer fewer than 0 or more than 30 words (see columns 4, 5 and 6).

The initial OLS regression model (column 1) shows that the significantly negative coefficient of *quitters* indicates that subjects studied approximately 11 words fewer, on average, when paired with the partners of low perseverance, compared to the single condition ($p = .05$). The results from the Tobit model (column 4) indicate a decrease of about 15 words when paired with such partners, although this effect is nearly significant ($p = .08$). The other regression models (columns 2, 3, 5 and 6) consistently revealed a significantly negative effect from the similar-quitter condition. Specifically, the Tobit model incorporating demographic variables shows that subjects in the similar-quitter condition studied around 27 words fewer on average than those in the single condition (column 6).

Our analysis also highlights several intriguing observations. First, the coefficient for high-quitter was not significant, indicating that higher-skilled partners did not negatively affect subjects' perseverance even when the partners exhibited less perseverance. Second, we found no evidence of

significant positive peer effects from high perseverance partners, indicating that higher perseverance did not translate into positive peer effects. Additionally, the variable *age* had positive impacts while *known_words* had negative impacts on the number of studied words (columns 3 and 5 at the top of Table 3).

Performance. We also conducted the same regression analysis regarding the test scores (the bottom of Table 3). Most of the main findings were similar to the results of former analysis regarding the number of studied words. The significant negative coefficient for *quitters* indicates that subjects' test scores decreased by approximately 1.8 words on average when they were paired with the partners possessing low perseverance (columns 1 and 4). The subjects in the similar-quitter condition consistently exhibited a negative effect across the four regressions. In this condition, the average test score was about 3 words lower compared to the single condition (columns 2, 3, 5 and 6 at the bottom of Table 3).

We replicated the analysis utilizing ANOVA (see Appendix A) and applied both Generalized Linear Models (GLM) with Poisson and Negative Binomial distributions as a robustness check.[8] The overall findings are consistent with those from the main regression analysis.

[8] We also utilized zero-truncated models due to the absence of zero counts in our dependent variables. The results from all GLMs were consistent with those from our main analysis and are available upon request.

Table 3

Regression Results (Study 1): The Effect of Conditions on Perseverance (Top) and Performance (Bottom)

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
quitters	-10.965*			-15.399		
	(5.522)			(8.761)		
finishers	1.110			0.723		
	(5.539)			(8.857)		
similar-quitter		-17.033**	-17.821**		-25.589*	-27.028**
		(6.366)	(6.236)		(10.044)	(9.779)
high-quitter		-5.281	-4.307		-5.733	-4.201
		(6.268)	(6.146)		(9.893)	(9.627)
similar-finisher		0.972	0.660		0.649	0.560
		(6.268)	(6.182)		(10.009)	(9.796)
high-finisher		1.263	0.329		0.811	-0.866
		(6.431)	(6.324)		(10.256)	(10.015)
age			0.488**			0.733*
			(0.182)			(0.289)
female			-5.661			-10.472
			(3.889)			(6.129)
college			-5.143			-7.671
			(4.339)			(6.849)
known_words			-0.087**			-0.126**
			(0.027)			(0.043)
Constant	42.357***	42.357***	37.209***	50.063***	50.001***	42.832**
	(4.574)	(4.564)	(8.917)	(7.313)	(7.270)	(14.121)
N	373	373	373	373	373	373

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
quitters	-1.756*			-1.814*		
	(0.887)			(0.895)		
finishers	-0.006			-0.045		
	(0.890)			(0.898)		
similar-quitter		-2.924**	-2.891**		-2.996**	-2.960**
		(1.017)	(1.014)		(1.023)	(1.014)
high-quitter		-0.662	-0.704		-0.703	-0.745
		(1.001)	(0.999)		(1.007)	(1.000)
similar-finisher		-0.788	-0.739		-0.830	-0.770
		(1.001)	(1.005)		(1.007)	(1.006)
high-finisher		0.865	0.754		0.830	0.718
		(1.027)	(1.028)		(1.034)	(1.029)
age			0.026			0.024
			(0.030)			(0.030)
female			-0.310			-0.321
			(0.632)			(0.632)
college			0.692			0.743
			(0.705)			(0.705)
known_words			0.008			0.009
			(0.004)			(0.004)
Constant	19.586***	19.586***	17.557***	19.659***	19.658***	17.643***
	(0.735)	(0.729)	(1.450)	(0.742)	(0.734)	(1.450)
N	373	373	373	373	373	373

Notes: The first three columns of the table report the results from ordinary least squares (OLS) regressions, while the last three columns present the outcomes of Tobit regressions. The base condition of these analyses is the Single condition. Standard errors in parentheses; *, ** and *** stand for statistical significance at the 5%, 1% and 0.1% level respectively.

Regression Analysis with IOS

In additional analysis, we examined the correlation between learning outcomes and IOS, describing the psychological distance between the subjects and their partners. We expected the positive correlation in cases where the partners possessed higher perseverance and a negative correlation when partners exhibited lower perseverance because larger peer effects were found from people with closer social distance in the previous literature (e.g., Bandiera et al., 2010).

We found a moderate level of negative correlation between the number of studied words and IOS when the partners possessed lower perseverance: r = -0.26 (p = .03) in the similar-quitter condition; r = -0.22 (p = .05) in the high-quitter condition. In the similar-finisher condition, we observed a moderate level of positive correlation (r = 0.21; p = .06), but not in the high-finisher condition (r = -0.15; p = .20).

We regressed the number of studied words on the treatment dummy variables (using the similar-quitter condition as the baseline treatment), IOS and their interaction variables. The results, presented in the top of Table 4, consistently show significant positive coefficients for the similar-finisher and IOS interaction term, indicating a more positive correlation between IOS and the number of studied words in the similar-finisher condition compared to the similar-quitter condition. These results held across four regression models: without control variables (1st column), with control variables (2nd column), Tobit regression without control variables (3rd column), and Tobit regression with control variables (4th column).

The correlation patterns between the test scores and IOS were similar to those observed between the number of studied words and IOS.[9] The bottom of Table 4 shows the same regressions in terms of the test scores. Again, we found the significant positive coefficients of the similar-finisher and IOS interaction term in the regressions.

[9] The results are as follows: In the similar-quitter condition, r = -0.21 (p = .07); in the high-quitter condition, r = -0.14 (p = .20); in the similar-finisher condition, r = 0.15 (p = .19); and in the high-finisher condition, r = -0.11 (p = .37).

Table 4

Regression Results (Study 1): The Effect of Conditions and IOS on Perseverance (Top) and Performance (Bottom)

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit
high-quitter	9.487	7.099	24.339	21.742
	(14.452)	(14.208)	(22.440)	(21.882)
similar-finisher	-17.989	-24.337	-23.265	-32.710
	(14.422)	(14.161)	(22.781)	(22.196)
high-finisher	11.558	10.957	27.930	27.017
	(14.696)	(14.242)	(23.198)	(22.334)
IOS	-10.270	-12.309*	-10.475	-13.393
	(5.422)	(5.334)	(8.320)	(8.134)
high-quitter#IOS	0.952	3.455	-3.012	0.390
	(7.182)	(7.114)	(11.009)	(10.808)
similar-finisher#IOS	16.538**	19.635**	21.437*	26.243**
	(6.248)	(6.122)	(9.766)	(9.512)
high-finisher#IOS	5.173	5.730	1.605	2.346
	(6.530)	(6.364)	(10.154)	(9.832)
known_words		-0.081**		-0.117**
		(0.028)		(0.044)
Constant	44.338***	34.239*	44.070*	28.456
	(10.935)	(13.133)	(16.975)	(20.350)
Controls	No	Yes	No	Yes
N	303	303	303	303

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit
high-quitter	1.423	1.940	1.513	2.066
	(2.387)	(2.415)	(2.376)	(2.388)
similar-finisher	-2.531	-2.464	-2.561	-2.473
	(2.382)	(2.407)	(2.370)	(2.379)
high-finisher	2.286	2.616	2.348	2.691
	(2.427)	(2.421)	(2.416)	(2.393)
IOS	-1.493	-1.346	-1.493	-1.335
	(0.895)	(0.907)	(0.891)	(0.896)
high-quitter#IOS	0.422	0.107	0.390	0.054
	(1.186)	(1.209)	(1.180)	(1.196)
similar-finisher#IOS	2.195*	2.148*	2.218*	2.162*
	(1.032)	(1.041)	(1.027)	(1.029)
high-finisher#IOS	0.969	0.725	0.959	0.705
	(1.078)	(1.082)	(1.073)	(1.069)
known_words		0.008		0.008
		(0.005)		(0.005)
Constant	19.426***	16.975***	19.426***	16.940***
	(1.806)	(2.233)	(1.797)	(2.206)
Controls	No	Yes	No	Yes
N	303	303	303	303

Notes: The first two columns of the table report the results from ordinary least squares (OLS) regressions, while the last two columns present the outcomes of Tobit regressions. The base condition of these analyses is the similar-quitter condition. In the second and fourth columns, variables controlled for include age, gender (noted as female), and attainment of a college degree or higher. Standard errors in parentheses; *, ** and *** stand for statistical significance at the 5%, 1% and 0.1% level respectively.

Discussion

In summary, our findings decisively demonstrate a negative impact from partners with low levels of perseverance, particularly when the skill levels were similar to the subjects. This negative influence was not evident from higher-skilled partners. Contrarily, we found no evidence of the positive effects from partners with higher perseverance levels. We entertained the possibility that

the failure to identify the positive effects was due to a ceiling effect, suggesting that there was little room for already motivated subjects to increase their study efforts. Indeed, we found the average enjoyment on the experiment was pretty high (7.51 out of 10)

Our analysis regarding IOS implies that lower social proximity accounts for the absence of overall positive peer effects. Prior research has investigated the relationship between peer effects and social distance (Bandiera et al., 2010; Bicchieri et al., 2022). Specifically, Bicchieri et al. (2022) underscore the pivotal role of social proximity (e.g., higher IOS) on enhancing positive peer effects in norm compliance behaviors. They demonstrated that without social proximity, there's a decline in norm adherence; however, in the presence of social proximity, norm erosion is averted as subjects react to both compliant and non-compliant peers' behaviors. This implies that the solely negative peer effects observed in our primary analysis could stem from insufficient social proximity toward partners. This idea is further supported by our data showing that the mean IOS score in paired conditions was a mere 2.19 out of 7, indicating a low average level of felt social proximity to partners.

Moreover, our analysis suggested the significance of skill similarity of partners in fostering positive peer effects. Our data indicated some evidence of a positive association between IOS and the number of words studied in the similar-finisher condition. This suggests that having partners of similar skill levels, combined with social proximity, is pivotal for positive peer influence. This interpretation is corroborated by research showing the presence of positive peer effects among subjects of similar abilities (Booij et al., 2017; Carrell et al., 2013).

Study 2: Sample with Lower Motivation in English Vocabulary Learning

In Study 1, we identified negative peer effects from low perseverance partners, especially when the partner's skills were similar to those of the subjects. Despite this, no overarching positive peer

effects were detected. We considered the possibility that the failure to detect positive peer effects in Study 1 could be attributed to a ceiling effect, which implies that the subjects who were already motivated had little room to further increase their study efforts.

Therefore, in Study 2, we recruited subjects with a lower baseline of motivation to reexamine the peer effects, specifically to investigate the potential positive effects from high perseverance partners.

Subjects. We recruited 400 subjects for our experiment through Lancers. We excluded 3 subjects declaring to know more than 400 words, as in Study 1, because their vocabulary level was too high for this experiment. As a result, we obtained a demographically diverse sample of 397 subjects in total (43% female, M_{age} = 40.66 years, age range: 18-76 years). Of these subjects, 47% are fully employed, 15% hold part-time positions, 2% are students, and the remaining 37% fall into other categories. The participation requirements for this study were identical to those in Study 1, except that subjects must not have engaged in English language study for more than six months.[10] The subjects received a fixed fee of 600 yen if they completed all the parts of the experiment. The duration of the experiment varied depending on how many English words they wanted to learn, with a median duration of approximately 39 minutes.

Design and procedure. The procedure was identical to that of Study 1 except for one question. Right before the question asking how fun the study was, we added the intrinsic motivation scale (Schmidt et al., 1996). Incorporating this scale aimed to validate the participation of subjects with weak motivation to learn English. The scale was a 6-point scale (1: Strongly disagree – 6: Strongly

[10] Not engaging English study refers to not taking English lessons or classes, not studying independently with English materials or apps, not watching English videos or podcasts, not learning by watching English movies or TV dramas, not reading English books or blogs, and not placing oneself in an environment where English is spoken.

agree). The scale included four questions such as "I enjoy learning English very much".[11] We calculated the average of these responses to create the variable *motivation* ($\alpha = 0.74$).

Results and Discussion

Manipulation Check

The manipulations regarding the skills and perseverance of the partners in this study were successful. Subjects in the high-skilled condition believed their partner had significantly higher vocabulary levels (M = 5.51, SD = 1.41) compared to those in the similar-skilled partners (M = 4.07, SD = 1.12; $F_{(1,319)} = 101.26$, $p < .001$, $\eta^2 = .24$). Subjects in the finisher conditions perceived that their partner studied more (M = 6.70, SD = 2.38) compared to those in the quitter conditions (M = 3.75, SD = 2.66; $F_{(1,319)} = 108.09$, $p < .001$, $\eta^2 = .25$).

Descriptive Analysis

Table 5 shows the selected characteristics of the subjects for each treatment including variables *known_words*, *eng_level*, *age* and *female* as in Study 1. The differences in these variables between the treatments were not statistically significant, indicating the randomization worked well (all p-values above 0.1).[12] The mean value of motivation was 2.90 (the minimum possible value is 1 and the maximum possible value is 6), which was less than the middle point of 3.5 (t (396) = 13.81; $p <$ 0.001), suggesting their motivation toward learning English was not high.

[11] We excluded the one question "I wish I could learn English in an easier way, without going to class" from the original list of Schmidt et al. (1996) because not all the subjects were students.

[12] As we expected, the degree of intrinsic motivation did not differ between treatments (ANOVA: $F_{(4, 392)} = 0.73$, $p = .57$), indicating the variable *motivation* described stable personality trait.

Table 5

Descriptive Statistics (Study 2)

	Similar-quitter	**High-quitter**	**Similar-finisher**	**High-finisher**	**single**	**Statistic test**	**p-value**
known_words	37.97	33.48	31.15	35.79	32.11	$F_{(4,392)} = 0.30$	$p = .88$
	(43.34)	(38.40)	(57.88)	(43.12)	(40.23)		
eng_level	1.38	1.51	1.25	1.31	1.38	$F_{(4,392)} = 1.94$	$p = .10$
	(0.65)	(0.71)	(0.55)	(0.57)	(0.67)		
motivation	2.98	2.88	2.78	2.87	2.98	$F_{(4,392)} = 0.73$	$p = .57$
	(0.98)	(0.88)	(0.66)	(0.87)	(0.89)		
age	41.22	39.37	40.40	41.30	41.28	$F_{(4,392)} = 0.53$	$p = .71$
	(11.42)	(10.42)	(9.3)	(12.29)	(10.03)		
female	0.39	0.39	0.50	0.44	0.43	$\chi_{(4)} = 2.65$	$p = .62$
N	78	94	72	77	76		

Notes: Standard deviations in parentheses. The variable *known_words* represents the number of words subjects knew in the pre-learning part; *eng_level* denotes their self-assessed level of English proficiency; *motivation* describes the average value of the intrinsic motivation scale (Schmidt et al., 1996); and *female* indicates the proportion of female subjects.

The characteristics of the sample were different between Studies 1 and 2 as expected (Table 6). The variable *known_words* was smaller in Study 2 (t (615) = 11.66; $p < 0.001$) and the level of English proficiency was lower in Study 2 (t (674) = 9.12; $p < 0.001$). The subjects in Study 2 enjoyed the experiment less (t (759) = 4.93; $p < 0.001$). These findings implied the subjects in Study 2 were less motivated to learn English.

Table 6

Summary Statistics in Studies 1 and 2

	Study 1	**Study 2**	**t-test**	**p-value**
known_words	84.60	34.13	11.66	$p < .001$
	(71.62)	(44.52)		
eng_level	1.89	1.38	9.12	$p < .001$
	(0.89)	(0.64)		
enjoy	7.51	6.70	4.93	$p < .001$
	(2.34)	(2.23)		
# of studied words	42.36	38.20	0.65	$p < .52$
(single treatment)	(39.37)	(37.72)		
# of correct words	19.59	17.14	2.45	$p < .02$
(single treatment)	(5.65)	(6.39)		

Notes: Standard deviations in parentheses. The variable *known_words* represents the number of words subjects knew in the pre-learning part; *eng_level* denotes their self-assessed level of English proficiency; and *enjoy* measures the overall enjoyment subjects reported for learning English words in the experiment.

Regression Analysis

Perseverance. We initially conducted our regression analysis using an OLS method to evaluate the impact of learning partners' perseverance on the number of words studied, following the same procedure as in Study 1 (refer to the top of Table 7). The models were identical to those used in Study 1, with the exception that *motivation* was included as a control variable in the models

incorporating demographic variables. And then, as in Study 1, we repeated these three analyses using Tobit regression methods (see columns 4, 5, and 6).

In column 1, the OLS model revealed a significantly negative coefficient for *quitters*, indicating that subjects studied approximately 10 words fewer on average when paired with low perseverance partners, compared to the single condition ($p = .04$). The Tobit model showed a decrease of about 15 words when paired with such partners ($p = .05$). This substantial overall negative effect on perseverance is similar to the findings from Study 1.

Interestingly, the other regression models consistently demonstrated a significant negative effect from the high-quitter condition rather than the similar-quitter one, diverging from the results of Study 1. Specifically, the Tobit model including demographic variables indicated that subjects in the high-quitter condition studied around 16 words fewer on average than those in the single condition (column 5). However, the coefficient for similar-quitter was not significant, indicating that similar-skilled partners did not negatively affect subjects' perseverance in Study 2.

As in Study 1, we found no evidence of significant positive peer effects from high perseverance partners. Additionally, the variable *age* and *motivation* had positive impacts on the number of studied words (columns 3 and 5).

Performance. We also conducted a regression analysis on the test scores (the bottom of Table 3). The main findings largely align with those from the analysis of the number of studied words. Specifically, subjects in the high-quitter condition consistently demonstrated a negative effect across the four regressions. In this condition, the average test score was about 1.8 words lower compared to the single condition (columns 2, 3, 5 and 6). However, the coefficient for *quitters* was negative but not significant (columns 1 and 4). We found no evidence of significant positive peer effects from high perseverance partners on own performance. Additionally, motivation positively influenced the test scores.

We replicated the analysis utilizing ANOVA (see Appendix A) and applied both Generalized Linear Models (GLM) with Poisson and Negative Binomial distributions as a robustness check.[13] The overall findings are consistent with those from the main regression analysis.

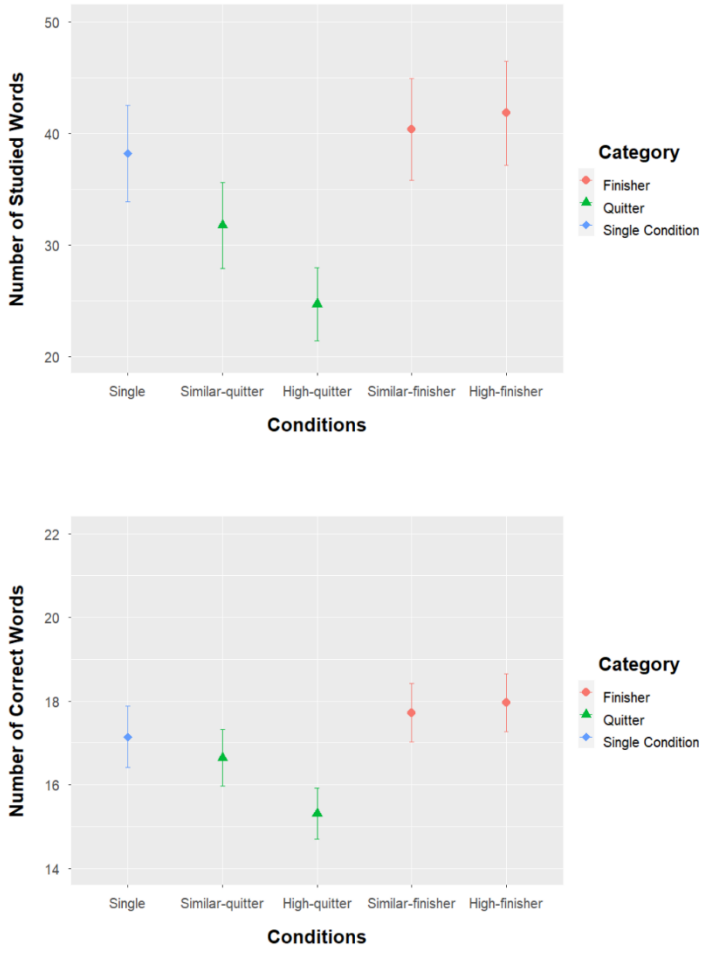

Figure 4. The Number of Studied Words in the Voluntarily Learning Part (Top) and the Number of Correct Words in the Test Part (Bottom) for each Treatment (Study 2).

Means are represented by three distinct shapes: a circle for Finisher, a triangle for Quitter, and a diamond for Single condition. Standard error bars are included.

[13] We also utilized zero-truncated models due to the absence of zero counts in our dependent variables. The results from all GLMs were consistent with those from our main analysis and are available upon request.

Table 7

Regression Results (Study 2): The Effect of Conditions on Perseverance (Top) and Performance (Bottom)

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
quitters	-10.325*			-14.690*		
	(5.035)			(7.425)		
finishers	2.903			3.448		
	(5.153)			(7.643)		
similar-quitter		-6.467	-6.302		-9.625	-9.049
		(5.895)	(5.783)		(8.655)	(8.426)
high-quitter		-13.527*	-12.044*		-18.856*	-16.295*
		(5.642)	(5.554)		(8.285)	(8.083)
similar-finisher		2.150	3.637		2.232	4.728
		(6.015)	(5.892)		(8.886)	(8.639)
high-finisher		3.608	4.411		4.606	5.945
		(5.913)	(5.781)		(8.786)	(8.529)
age			0.539**			0.830***
			(0.168)			(0.247)
female			4.966			7.321
			(3.668)			(5.374)
college			3.066			4.055
			(3.781)			(5.538)
known_words			-0.061			-0.088
			(0.042)			(0.061)
motivation			6.739**			10.004**
			(2.119)			(3.143)
Constant	38.197***	38.197***	-5.948	43.885***	43.866***	-22.937
	(4.193)	(4.195)	(10.402)	(6.212)	(6.198)	(15.367)
N	397	397	397	397	397	397

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
quitters	-1.232			-1.208		
	(0.833)			(0.838)		
finishers	0.701			0.704		
	(0.853)			(0.857)		
similar-quitter		-0.504	-0.824		-0.476	-0.799
		(0.975)	(0.946)		(0.978)	(0.943)
high-quitter		-1.836*	-1.899*		-1.815	-1.875*
		(0.933)	(0.908)		(0.936)	(0.905)
similar-finisher		0.577	0.927		0.553	0.906
		(0.995)	(0.964)		(0.998)	(0.961)
high-finisher		0.816	0.851		0.846	0.877
		(0.978)	(0.945)		(0.981)	(0.942)
age			0.053			0.052
			(0.027)			(0.027)
female			-0.068			-0.086
			(0.600)			(0.598)
college			1.688**			1.681**
			(0.618)			(0.617)
known_words			0.010			0.010
			(0.007)			(0.007)
motivation			1.324***			1.337***
			(0.347)			(0.346)
Constant		17.145***	9.823***		17.145***	9.847***
		(0.694)	(1.701)		(0.696)	(1.696)
N		397	397		397	397

Notes: The first three columns of the table report the results from ordinary least squares (OLS) regressions, while the last three columns present the outcomes of Tobit regressions. The base condition of these analyses is the Single condition. Standard errors in parentheses; *, ** and *** stand for statistical significance at the 5%, 1% and 0.1% level respectively.

Regression Analysis with Motivation

Next, we focused our analysis on motivation. The variable *motivation* was positively correlated with the number of studied words ($r = 0.15$; $p = 0.003$) and the test scores ($r = 0.20$; $p < 0.001$).

To probe further, we divided the sample into two groups based on the median motivation value of 3. Initially, we conducted a Tobit regression analysis on the number of studied words, using treatment dummy variables, demographic variables, and *known_words* as the explanatory variables (Table 8). Among the less motivated sample (column 1), the coefficients for *high-quitter* and *similar-quitter* were significantly negative. Contrarily, among the more motivated sample (column 2), they were not statistically significant. This result suggests that negative peer effects primarily emanate from the less motivated group.

A parallel approach was employed to analyze the test scores (columns 3 and 4), and we found similar results with the one regarding perseverance. Among the less motivated sample, the coefficient for *high-quitter* was significantly negative ($p < 0.001$), although the coefficient for *similar-quitter* was not statistically significant. Among the more motivated sample, the coefficients for both *high-quitter* and *similar-quitter* were not statistically significant.

Regression Analysis with IOS

In alignment with the approach in Study 1, we examined the relationship between learning outcomes and social proximity measured by IOS scale. Consistent with our expectations, we found a moderate negative correlation between the number of studied words and IOS when the partners possessed lower perseverance: $r = -0.17$ ($p = .13$) in the similar-quitter condition; $r = -0.02$ ($p = .88$) in the high-quitter condition. Furthermore, we found a moderate positive correlation between the number of studied words and IOS when the partners possessed higher perseverance: $r = 0.23$ ($p = .05$) in the similar-finisher condition; $r = 0.12$ ($p = .28$) in the high-finisher condition, aligning with our hypothesis as well. A similar correlation was observed between the test scores and IOS.[14]

[14] The results are as follows: $r = -0.13$ ($p = .28$) in the similar-quitter condition; $r = 0.11$ ($p = .31$) in the high-quitter condition; $r = 0.27$ ($p = .02$) in the similar-finisher condition; and $r = 0.13$ ($p = .25$) in the high-finisher condition

Table 8

Tobit Regression Results by Motivation-divided Group (Study 2): The Effect of Conditions on Perseverance and Performance

	(1)	(2)	(3)	(4)
	DV = Perseverance Less motivated	DV = Perseverance More motivated	DV = Performance Less motivated	DV = Performance More motivated
similar-quitter	-32.992**	6.982	-1.461	-0.335
	(11.924)	(11.928)	(1.395)	(1.294)
high-quitter	-42.146***	2.777	-3.436*	-0.385
	(11.292)	(11.666)	(1.313)	(1.272)
similar-finisher	-12.693	12.802	0.628	0.784
	(11.684)	(12.937)	(1.366)	(1.388)
high-finisher	-16.202	21.290	1.315	0.351
	(12.149)	(11.918)	(1.405)	(1.278)
age	0.819*	0.966**	0.043	0.073
	(0.327)	(0.368)	(0.039)	(0.039)
female	-3.239	19.100*	-0.689	0.725
	(7.259)	(7.919)	(0.857)	(0.847)
college	12.221	-5.397	1.850*	1.472
	(7.475)	(8.120)	(0.880)	(0.874)
known_words	0.042	-0.095	0.025	0.010
	(0.114)	(0.074)	(0.013)	(0.008)
Constant	18.618	-7.640	13.595***	13.054***
	(17.074)	(18.299)	(2.006)	(1.975)
N	193	204	193	204

Notes: Standard errors in parentheses; *, ** and *** stand for statistical significance at the 5%, 1% and 0.1% level respectively. Columns 1 and 3 include only the less intrinsically motivated sample (*motivation* < 3), while columns 2 and 4 contain only the more motivated sample (*motivation* ≥ 3). The variable *known_words* represents the number of words subjects knew in the pre-learning part; and *college* denotes if the subject has a college degree.

To further investigate these relationships, we regressed the number of studied words on the treatment dummy variables (with the similar-quitter condition as the baseline treatment), IOS and their interaction variables. The structure of these four regressions, presented at the top of Table 9,

paralleled our previous analyses. We consistently found significant positive coefficients for the interaction term between *similar-finisher* and IOS across all four regressions: without control variables (column 1), with control variables (column 2), Tobit regression without control variables (column 3), and Tobit regression with control variables (column 4). This suggests stronger positive correlation between IOS and the number of studied words in the similar-finisher condition, compared to the similar-quitter condition across the four regressions. The bottom of Table 9 displayed analogous regressions for the test scores. Again, we found the significant positive coefficients for the interaction term between *similar-finisher* and IOS.

Discussion

Study 2 successfully recruited subjects who were less motivated and less skilled. In line with the results of Study 1, Study 2 consistently demonstrated an overall negative impact from partners with lower levels of perseverance, and failed to find evidence supporting a positive impact from partners with higher levels of perseverance. Additionally, as in Study 1, we observed a positive association between IOS and the learning outcomes in the similar-finisher condition.

Interestingly, the negative effects in this study were observed from partners who were higher-skilled, diverging from Study 1, where the negative effect emerged from partners with similar skill levels. In addition, our nuanced analysis revealed that the subjects with lower intrinsic motivation towards English learning were primarily affected by these negative effects.

Table 9

Regression Results (Study 2): The Effect of Conditions and IOS on Perseverance (Top) and Performance (Bottom)

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit
high-quitter	-15.658	-14.582	-22.635	-20.613
	(10.663)	(10.576)	(15.340)	(15.097)
similar-finisher	-16.284	-15.082	-25.178	-22.917
	(11.533)	(11.510)	(16.897)	(16.734)
high-finisher	-7.990	-7.254	-13.723	-12.487
	(11.279)	(11.245)	(16.716)	(16.573)
IOS	-4.551	-4.635	-5.960	-5.995
	(3.194)	(3.162)	(4.586)	(4.507)
high-quitter#IOS	4.081	3.838	6.541	5.994
	(4.694)	(4.654)	(6.720)	(6.611)
similar-finisher#IOS	11.503*	10.898*	17.204*	16.115*
	(4.603)	(4.572)	(6.835)	(6.734)
high-finisher#IOS	8.175	7.788	12.496	11.852
	(4.358)	(4.332)	(6.494)	(6.415)
known_words		-0.044		-0.061
		(0.045)		(0.065)
Constant	41.182***	18.615	46.551***	12.448
	(7.790)	(11.307)	(11.237)	(16.347)
Controls	No	Yes	No	Yes
N	321	303	303	303

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit
high-quitter	-3.582*	-3.486*	-3.600*	-3.501*
	(1.749)	(1.719)	(1.749)	(1.707)
similar-finisher	-2.869	-2.359	-2.964	-2.448
	(1.892)	(1.871)	(1.891)	(1.858)
high-finisher	-1.265	-0.409	-1.255	-0.407
	(1.851)	(1.828)	(1.850)	(1.816)
IOS	-0.579	-0.599	-0.597	-0.614
	(0.524)	(0.514)	(0.524)	(0.510)
high-quitter#IOS	1.154	1.189	1.158	1.191
	(0.770)	(0.756)	(0.770)	(0.751)
similar-finisher#IOS	1.815*	1.779*	1.836*	1.798*
	(0.755)	(0.743)	(0.754)	(0.738)
high-finisher#IOS	1.161	0.888	1.159	0.889
	(0.715)	(0.704)	(0.714)	(0.699)
known_words		0.018*		0.018*
		(0.007)		(0.007)
Constant	17.844***	13.950***	17.908***	14.070***
	(1.278)	(1.838)	(1.278)	(1.826)
Controls	No	Yes	No	Yes
N	321	321	321	321

Notes: The first two columns of the table report the results from ordinary least squares (OLS) regressions, while the last two columns present the outcomes of Tobit regressions. The base condition of these analyses is the similar-quitter condition. In the second and fourth columns, variables controlled for include age, gender (noted as female), and attainment of a college degree or higher. Standard errors in parentheses; *, ** and *** stand for statistical significance at the 5%, 1% and 0.1% level respectively.

General Discussion and Conclusions

We conducted two experiments, both resembling a typical online English learning environment. These experiments scrutinize the clear causal effects of learning partners' perseverance on individuals' perseverance as well as performance, and how these effects varied with the partners'

skill levels, either being higher or similar to the subjects. Our sample, consisting of a diverse demographic of Japanese individuals, along with the realistic learning environments used in the studies, provides substantial external validity.

In Study 1, the subjects were actively engaged in studying English. In Study 2, we retained the experimental structure but enlisted subjects who had refrained from English study for a minimum of six months. A consistent finding across both studies was the substantial overall negative impact of partners characterized by lower perseverance levels. The number of words studied dropped by approximately 11 to 15 words. In Study 1, there was also a decrease of about 1.8 points in test scores. We did not encounter evidence of positive impacts from partners with higher levels of perseverance.

In Study 1, this negative influence was attributed to partners of similar skill levels, whereas in Study 2, it emanated from higher-skilled partners.

Deepening our analysis in Study 2 to examine the intrinsic motivation towards English learning, we discerned that the less intrinsically motivated sample was the primary contributor to these negative effects. This observation aligns with Tanaka's (2017) survey study, which focused on students with low intrinsic motivation. She identified the presence of negative peer effects, as well as the absence of positive peer effects, on one’s motivation in vocabulary learning.

Intriguingly, both studies provided some evidence suggesting that social proximity, measured by IOS, could foster positive peer effects when partners possess similar skills and higher perseverance levels. The results align with earlier research demonstrating pronounced peer effects among individuals sharing closer social distance (Bandiera et al., 2010).

Pronounced Negative Peer Effects and Absence of positive peer effects

Our findings demonstrate an asymmetry in positive and negative peer effects, where negative influences from peers are more potent. The identification of negative peer effects on individual

perseverance (e.g., Battaglini et al., 2017; Isler & Gächter, 2022; Thöni & Gächter, 2015), and the absence of positive influences (e.g., Bicchieri et al., 2022; Georganas et al., 2015; Isler & Gächter, 2022; Rauhut, 2013; Thöni & Gächter, 2015), are well-documented phenomena.

Particularly, Thöni and Gächter (2015) observed patterns similar to ours. They found that when subjects observed a peer exerting less (hypothetical) effort given a certain wage, they substantially decreased their own effort in a gift-exchange experiment in the lab. However, observing a peer choosing more effort did not consistently lead to an increase in their own effort. Isler and Gächter (2022)'s study on conforming behaviors suggested that negative peer effects may arise when a peer unexpectedly violates a norm.

This interpretation of Isler and Gächter (2022) could also explain the observed asymmetry in our study. Subjects in the single condition studied a significantly higher number of words on average compared to those whose partners exhibited low perseverance. Presumably expecting their partners to study a comparable amount, subjects were markedly affected when faced with their partners' unexpectedly low engagement. This discrepancy led to a more pronounced impact of negative effects compared to positive ones.

Sources of Negative Peer Effects in Studies 1 and 2

Although we found overall negative peer effects in both studies, the effect varied with the peer's skill level. Specifically, in Study 1, subjects were influenced by the similar-quitter treatment but not the high-quitter treatment. Conversely, in Study 2, they were affected by the high-quitter treatment but not the similar-quitter treatment. This section speculates on the reasons for these differences by drawing comparisons with existing literature.

Negative effects from partners with similar skill levels were observed in Study 1, possibly due to perceived lower self-efficacy when partners discontinued their efforts. This interpretation aligns with Schunk and DiBenedetto (2021), who noted that witnessing a similarly skilled peer fail can

reduce an individual's self-efficacy. Schunk and Hanson (1985) claimed that self-efficacy is critical in learning environments, influencing effort investment, persistence, and overall success.

Additionally, social learning theory underscores the tendency of individuals to mimic the behaviors of similar peers (Hanna et al., 2013). Also, the peer effect economic model proposed by Battaglini et al. (2005) suggests that an individual's perception of their own perseverance is negatively affected when they observe similar others failing to handle temptations.

Another interpretation from Study 1 suggests that subjects, already possessing a certain level of English knowledge, may have perceived that high-quitters—those with even higher ability than themselves—stopped learning because they already had substantial knowledge. Drawing upon the findings of Isler and Gächter (2022), it seems plausible that negative peer effects failed to occur, because the act of quitting was not sufficiently surprising to the subjects.

In Study 2, negative effects emerged among subjects paired with high-skilled partners, who were less proficient than those in Study 1. This pattern is consistent with previous research, which suggests that less successful individuals are more inclined to rely on social information and emulate the behaviors of peers who are successful, attractive, prestigious, or of high status (Atkisson et al., 2012; Hanna et al., 2013; Henrich & Gil-White, 2001; Mesoudi, 2008). Individuals are more likely to pay attention to and learn from those who are skilled at a task (Bandura, 1986), and to copy their strategies (Mesoudi & O'Brien, 2008; Mesoudi, 2008). Additionally, another research also shows negative influences from high-skilled peers among lower-ability individuals (Feld & Zölitz, 2017).

Positive Perception Bias

Furthermore, our auxiliary analysis of the perception of partners' influences on perseverance, obtained from the survey question, suggests that perceived influences does not strongly correlate with subjects' actual perseverance. In other words, subjects did not accurately estimate their partners' influences on the number of words they studied (see Appendix B for the detailed analysis).

Interestingly, the results also imply a *positive perception bias.* Specifically, among subjects paired with low perseverance peers, the majority believed they received no influence or even a positive influence from their partners, despite pronounced negative effects actually being observed across both studies. Conversely, among subjects paired with high perseverance partners, the majority believed they received positive influences, although non-significant positive effects were actually observed in our studies.

These findings imply that perception measures, commonly used in previous studies (e.g., Keefe, 1994; Steinberg & Silverberg, 1986; Urberg et al., 1990), do not accurately reflect subjects' actual behaviors and highlight the importance of directly measuring behaviors.

Implications

Our research primarily focused on the peer effects in online English learning due to the ease of implementation. However, our findings can be applied broadly, thanks to the vast array of online courses available in diverse domains, such as programming, art, science, business management, and even online fitness or mindfulness courses like yoga and meditation. The impact of peer effects in these various fields may diverge from our current findings, offering deeper insights into the nature of peer influences.

Furthermore, this study has several other potential applications, including exploring cross-cultural peer effects and testing various interventions, such as setting default options to continue studying or sharing more detailed information about learning partners, like their political preferences or favorite movies. Our experimental design, implemented via Qualtrics, can facilitate such studies, and the file is available upon request.

To address the issue of attrition, online learning platforms have proactively developed social features to promote user engagement. Our study provides two practical suggestions for enhancing user-matching algorithms to optimize motivation and engagement. First, caution is advised when

pairing users with partners who display low perseverance, especially for users who are less intrinsically motivated and hence more susceptible to negative peer influences. Second, matching users with socially proximate partners who are similarly skilled and display higher perseverance levels can increase engagement and motivation. These user-matching strategies have the potential to be effective not only in online educational settings but also in various other collaborative scenarios.

Acknowledgements: The authors thank Satoshi Akutsu, Yukihiko Funaki, Noboyuki Hanaki, Yoshio Kamijyo, Daniel-Navarro Martinez, Charles Noussair, Erica Okada, Jordi Quoidbach, Tiffany Tsz Kwan Tse, Marie-Claire Villeval, Marc Willinger, and Takeshi Nishimura for helpful comments and suggestions. Special thanks to Koki Shimazu for his support in creating 500 English vocabulary questions for the experiments and to Ryan Wong, a research assistant, for his help in creating graphs.

Competing Interests Statement: The authors declare that they have no conflict of interest. The experiment reported in this paper has been approved by the ethical committee at Hitotsubashi University (No. 2021A011).

Financial Support Statement: This research has benefited from the financial support of (a) Grants-in-aid for Scientific Research No. 21K13257 from the Japan Society for the Promotion of Science, and (b) Sasakawa Scientific Research Grant No. 2022-1032 from the Japan Science Society.

Supplementary Materials: All the experimental instructions and raw data from the entire experiments are available on the OSF website (https://osf.io/7qc35/).

References

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98–108. https://doi.org/10.1016/j.labeco.2014.05.008

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology*, *10*(3), 519–537. https://doi.org/10.1177/147470491201000309

Bandiera, O., Barankay, I., & Rasul, I. (2010). Social incentives in the workplace. *Review of Economic Studies*, *77*(2), 417–458. https://doi.org/10.1111/j.1467-937X.2009.00574.x

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* (pp. xiii, 617). Prentice-Hall, Inc.

Battaglini, M., Bénabou, R., & Tirole, J. (2005). Self-control in peer groups. *Journal of Economic Theory*, *123*(2), 105–134. https://doi.org/10.1016/j.jet.2005.04.001

Battaglini, M., Díaz, C., & Patacchini, E. (2017). Self-control and peer groups: An empirical analysis. *Journal of Economic Behavior & Organization*, *134*, 240–254. https://doi.org/10.1016/j.jebo.2016.12.018

Beugnot, J., Fortin, B., Lacroix, G., & Villeval, M. C. (2019). Gender and peer effects on performance in social networks. *European Economic Review*, *113*, 207–224. https://doi.org/10.1016/j.euroecorev.2019.01.002

Bicchieri, C., Dimant, E., Gächter, S., & Nosenzo, D. (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, *132*, 59–72. https://doi.org/10.1016/j.geb.2021.11.012

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *The Review of Economic Studies*, *84*(2), 547–578. https://doi.org/10.1093/restud/rdw045

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.

Buechel, B., Mechtenberg, L., & Petersen, J. (2014). Peer effects and students' self-control. *MPRA Paper*, Article 53658. https://ideas.repec.org//p/pra/mprapa/53658.html

Buechel, B., Mechtenberg, L., & Petersen, J. (2018). If I can do it, so can you! Peer effects on perseverance. *Journal of Economic Behavior & Organization*, *155*, 301–314. https://doi.org/10.1016/j.jebo.2018.09.004

Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, *9*(1), 131–153. https://doi.org/10.1146/annurev-economics-063016-103625

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, *27*(3), 439–464. https://doi.org/10.1086/600143

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, *81*(3), 855–882. https://doi.org/10.3982/ECTA10168

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739–1774. https://doi.org/10.1257/aer.101.5.1739

Eriksson, T., Adawi, T., & Stöhr, C. (2017). "Time is the bottleneck": A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*, *29*(1), 133–146. https://doi.org/10.1007/s12528-016-9127-8

Feld, J., & Zölitz, U. (2017). Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, *35*(2), 387–428. https://doi.org/10.1086/689472

Georganas, S., Tonin, M., & Vlassopoulos, M. (2015). Peer pressure and productivity: The role of observing and being observed. *Journal of Economic Behavior & Organization*, *117*, 223–232. https://doi.org/10.1016/j.jebo.2015.06.014

Gerhards, L., & Gravert, C. (2020). Because of you I did not give up – Peer effects in perseverance. *Journal of Economic Psychology*, *81*, 102316. https://doi.org/10.1016/j.joep.2020.102316

Golsteyn, B. H. H., Non, A., & Zölitz, U. (2021). The impact of peer personality on academic achievement. *Journal of Political Economy*, *129*(4), 1052–1099. https://doi.org/10.1086/712638

Hanna, R. C., Crittenden, V. L., & Crittenden, W. F. (2013). Social learning theory: A multicultural study of influences on ethical behavior. *Journal of Marketing Education*, *35*(1), 18–25. https://doi.org/10.1177/0273475312474279

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196. https://doi.org/10.1016/S1090-5138(00)00071-4

Isler, O., & Gächter, S. (2022). Conforming with peers in honesty and cooperation. *Journal of Economic Behavior & Organization*, *195*, 75–86. https://doi.org/10.1016/j.jebo.2021.12.026

Jaccard, J., Blanton, H., & Dodge, T. (2005). Peer influences on risk behavior: An analysis of the effects of a close friend. *Developmental Psychology*, *41*(1), 135–147. https://doi.org/10.1037/0012-1649.41.1.135

Keefe, K. (1994). Perceptions of normative social pressure and attitudes toward alcohol use: Changes during adolescence. *Journal of Studies on Alcohol*, *55*(1), 46–54. https://doi.org/10.15288/jsa.1994.55.46

Kruger, J., & Gilovich, T. (1999). "Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology*, *76*(5), 743–753. https://doi.org/10.1037/0022-3514.76.5.743

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204. https://doi.org/10.1016/j.compedu.2004.12.004

Lomi, A., Snijders, T. A. B., Steglich, C. E. G., & Torlo, V. J. (2011). Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance. *Social Science Research*, *40*(6), 1506–1520. https://doi.org/10.1016/j.ssresearch.2011.06.010

Lyle, D. S. (2007). Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point. *The Review of Economics and Statistics*, *89*(2), 289–299.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542. https://doi.org/10.2307/2298123

Mas, A., & Moretti, E. (2009). Peers at work. *American Economic Review*, *99*(1), 112–145. https://doi.org/10.1257/aer.99.1.112

Mesoudi, A. (2008). An experimental simulation of the "copy-successful-individuals" cultural learning strategy: Adaptive landscapes, producer–scrounger dynamics, and informational

access costs. *Evolution and Human Behavior*, *29*(5), 350–363. https://doi.org/10.1016/j.evolhumbehav.2008.04.005

Mesoudi, A., & O'Brien, M. J. (2008). The cultural transmission of great basin projectile-point technology I: An experimental simulation. *American Antiquity*, *73*(1), 3–28. https://doi.org/10.1017/S0002731600041263

Narayanasamy, S. K., & Elçi, A. (2020). An effective prediction model for online course dropout rate. *International Journal of Distance Education Technologies (IJDET)*, *18*(4), 94–110. https://doi.org/10.4018/IJDET.2020100106

Nation, I. S. P. (Ed.). (2001). The Cambridge Applied Linguistics Series. In *Learning vocabulary in another language* (3rd ed., pp. ii–iv). Cambridge University Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Rauhut, H. (2013). Beliefs about lying and spreading of dishonesty: Undetected lies and their constructive and destructive social dynamics in dice experiments. *PLOS ONE*, *8*(11), e77878. https://doi.org/10.1371/journal.pone.0077878

Rosaz, J., Slonim, R., & Villeval, M. C. (2016). Quitting and peer effects at work. *Labour Economics*, *39*, 55–67. https://doi.org/10.1016/j.labeco.2016.02.002

Ross, L., & Ward, A. (1996). Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding. In *Values and Knowledge*. Psychology Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61,* 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics, 6*(1), 253–272. https://doi.org/10.1146/annurev-economics-071813-104217

Schmidt, R., Boraie, D., & Kassabgy, O. (1996). Foreign Language motivation: Internal structure and external connections. In *Language learning motivation: Pathways to the new century* (pp. 9–70). University of Hawaii, Second Language Teaching & Curriculum Center. http://hdl.handle.net/10125/40769

Schunk, D. H., & DiBenedetto, M. K. (2021). Chapter four—Self-efficacy and human motivation. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 8, pp. 153–179). Elsevier. https://doi.org/10.1016/bs.adms.2020.10.001

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*(3), 313–322. https://doi.org/10.1037/0022-0663.77.3.313

Shimazu, K. (2021). *Bijuaru de oboeru IELTS kihon eitango*. Japan taimuzu shuppan.

Steinberg, L., & Silverberg, S. B. (1986). The vicissitudes of autonomy in early adolescence. *Child Development, 57*(4), 841–851. https://doi.org/10.2307/1130361

Suls, J., Wan, C. K., & Sanders, G. S. (1988). False consensus and false uniqueness in estimating the prevalence of health-protective behaviors. *Journal of Applied Social Psychology, 18*(1), 66–79. https://doi.org/10.1111/j.1559-1816.1988.tb00006.x

Tanaka, M. (2017). Examining EFL vocabulary learning motivation in a demotivating learning environment. *System, 65,* 130–138. https://doi.org/10.1016/j.system.2017.01.010

Tello, S. F. (2007). An analysis of student persistence in online education: *International Journal of Information and Communication Technology Education*, *3*(3), 47–62. https://doi.org/10.4018/jicte.2007070105

Thöni, C., & Gächter, S. (2015). Peer effects and social preferences in voluntary cooperation: A theoretical and experimental analysis. *Journal of Economic Psychology*, *48*, 72–88. https://doi.org/10.1016/j.joep.2015.03.001

Urberg, K. A., Shyu, S. J., & Liang, J. (1990). Peer influence in adolescent cigarette smoking. *Addictive Behaviors*, *15*(3), 247–255. https://doi.org/10.1016/0306-4603(90)90067-8

Van Boven, L., Kamada, A., & Gilovich, T. (1999). The perceiver as perceived: Everyday intuitions about the correspondence bias. *Journal of Personality and Social Psychology*, *77*(6), 1188–1199. https://doi.org/10.1037/0022-3514.77.6.1188

Vantage Market Research. (2023). *Online Education/E-Learning Market—Global industry assessment & forcast*. Vantage Market Research. https://www.vantagemarketresearch.com

Vorauer, J. D., & Miller, D. T. (1997). Failure to recognize the effect of implicit social influence on the presentation of self. *Journal of Personality and Social Psychology*, *73*(2), 281–295. https://doi.org/10.1037/0022-3514.73.2.281

Zabell, S. (2023, June 6). *4 reasons Duolingo is more fun with friends*. Duolingo Blog. https://blog.duolingo.com/friends-social-features/

Appendix A: Analysis with ANONA

For the main analysis, we employed regression analysis. In this section, we replicate the analysis utilizing ANOVA. It is important to note that, unlike Tobit regressions, ANOVA does not account for the constraints that subjects could not study fewer than 1 or more than 95 words, or correctly answer fewer than 0 or more than 30 words. Nonetheless, the overall patterns are consistent with the main regression analysis, indicating that our results are robust.

Study 1

Perseverance. The main effect of experimental condition on the number of their studied words was significant (analysis of variance (ANOVA): $F_{(4,368)} = 3.04$, $p = .02$, $\eta^2 = .03$). We further conducted multiple comparisons with Holm's sequentially rejective Bonferroni procedure. As shown in Figure 3, there were negative effects from the partners with low perseverance. Subjects in the similar-quitter condition ($M = 25.32$, $SD = 32.13$) studied fewer words compared to those in the single condition ($M = 42.36$, $SD = 39.37$), indicating the negative peer effect (t (368) = 2.68, $p = .05$). The number of studied words in the similar-quitter condition was also significantly lower than the similar-finisher ($M = 43.33$, $SD = 40.62$; t (368) = 2.91, $p = .04$) and high-finisher conditions ($M = 43.62$, $SD = 40.45$; t (368) = 2.88, $p = .04$). There was no statistically significant difference between those in single and high-quitter condition ($M = 37.08$, $SD = 37.72$; t (368) = 0.84, $p > .99$). On the other hand, positive effects stemming from peers with high perseverance were not found; the number in the single condition was not different from that in either the similar-finisher (t (368) = 0.16, $p > .99$) or high-finisher condition (t (368) = 0.20, $p > .99$).

Performance. The results of ANOVA revealed the statistically significant main effect of experimental condition on the test scores ($F_{(4,368)} = 3.86$, $p = .004$, $\eta^2 = .04$). Negative peer effects were observed, while positive peer effects were not. Specifically, the similar-quitter condition yielded a lower average test scores ($M = 16.66$, $SD = 5.71$) compared to the single condition ($M =$

19.59, *SD* = 5.65; *t* (368) = 2.88, *p* = .03), although no significant differences were found between the high-quitter (*M* = 18.92, *SD* = 6.66) and single conditions (*t* (368) = 0.66, *p* > .99). This pattern of results confirms that negative peer effects emanate exclusively from similarly skilled partners. Further, the absence of positive peer effects was confirmed, as the test scores in the similar-finisher condition (*M* = 18.80, *SD* = 6.39) and in the high-finisher condition (*M* = 20.45, *SD* = 5.92) did not statistically differ from the single condition (*t* (368) = 0.79, *p* > .99; *t* (368) = 0.84, *p* > .99)

Study 2

Perseverance. We found a statistically significant main effect of the conditions on the number of studied words was significant (analysis of variance (ANOVA): $F_{(4, 392)} = 3.22$, $p = .01$, $\eta^2 = .03$). We further conducted multiple comparisons with Holm's sequentially rejective Bonferroni procedure as in Study 1. As shown in Figure 4, there appears negative effects from the partners with low perseverance due to a noticeable difference of 13.53 words between that in high-quitter and single conditions. Nonetheless, the number in the high-quitter condition (*M* = 24.67, *SD* = 31.92) and that in the similar-quitter condition (*M* = 31.73, *SD* = 34.22) did not significantly differ from that in the single condition (*M* = 38.20, *SD* = 37.72; *t* (392) = 2.40, *p* = .10 and *t* (392) = 1.10, *p* = .99). Note that we did find that the number was significantly smaller in the high-quitter condition compared to the similar-finisher (*t* (392) = 2.74, *p* = .04) or high-finisher conditions (*t* (392) = 3.05, *p* = .02), indicating the negative impacts from the high-skilled partners quitting midway compared to finishers. On the other hand, no positive peer effects were found as in Study 1. The number of studied words in both the similar-finisher (*M* = 40.35, *SD* = 38.67) and the high-finisher condition (*M* = 41.81, *SD* = 40.79) was not different from the single condition (*t* (392) = 0.36, *p* = .99 and *t* (392) = 0.61, *p* = .99).

Performance. The results of ANOVA revealed the statistically significant main effect of experimental condition on test scores ($F_{(4, 392)} = 2.62$, $p = .03$). Figure 4 indicates negative effects,

characterized by a noticeable difference of 1.84 words, constituting approximately 6% of all questions, between the high-quitter and single conditions. Yet, the test scores in the high-quitter condition (M = 15.31, SD = 5.92) as well as that in the similar-quitter condition (M = 16.64, SD = 5.97) did not significantly differ from that in the single condition (M = 17.14, SD = 6.39; t (392) = 1.97, p = .30 and t (392) = 0.52, p = .99). Importantly, lower performance was indeed observed when the high-quitter condition was compared to the high-finisher condition because the test scores in the high-quitter condition was lower compared to that in the high-finisher condition (t (392) = 2.85, p = .05). As in Study 1, no positive peer effects were found; the test scores in both the similar-finisher (M = 17.72, SD = 5.89) and the high-finisher condition (M = 17.96, SD = 6.07) was not statistically different from the single condition (t (392) = 0.58, p = .99 and t (392) = 0.83, p = .99).

Appendix B: Perception of the Influences from Partners

Across both studies, we consistently found an overall negative impact from partners with lower perseverance. However, we did not find evidence of positive impacts from partners with higher levels of perseverance. This raises an important question: Did subjects accurately perceive the influences from their partners, or was their perception systematically biased? Addressing this question could offer deeper insights into the absence of overall positive peer effects.

According to the literature, common measures of peer effects, such as self-reports on the extent to which individuals believe their behavior results from peer effects (e.g., Keefe, 1994; Steinberg & Silverberg, 1986; Urberg et al., 1990), may not be reliable. This unreliability stems from individuals' inability to accurately evaluate the influences of their peers. A related concept is *naïve realism* (Ross & Ward, 1996), where people tend to believe they see the world exactly as it is, without bias. They think their attitudes and beliefs arise from an unbiased view of facts, yet they often underestimate their own biases (Kruger & Gilovich, 1999; Nisbett & Wilson, 1977; Van Boven et al., 1999).

Individuals tend to underestimate the influences from peers. For example, subjects failed to perceive the effect of peers' self-descriptions on their own self-descriptions, while observers indicated that they were indeed influenced by peers' self-descriptions (Vorauer & Miller, 1997). However, individuals may also overestimate the extent of peer effects to justify their decisions and past behavior (Jaccard et al., 2005; Suls et al., 1988).

Therefore, we conducted an analysis to explore if the perception of the influences from partners aligns with the actual results of perseverance. To the best of our knowledge, no studies have examined such perception bias in both positive and negative peer effects.

Study 1

Surprisingly, we did not find a significant correlation between the number of studied words and the perception of the influences except in the similar-finisher condition ($r = .058$, $p = .62$ in the similar-quitter; $r = .095$, $p = .41$ in the high-quitter; $r = .43$, $p < .001$ in the similar-finisher; $r = .14$, $p = .23$ in the high-finisher). This indicates that the perception of the partners' influences, especially negative influences, does not strongly reflect the actual perseverance.

From the 7-point scale question regarding the perception of the partner's influence on the amount of words learned by subjects, we created the variable *partner_influence*. This variable takes the value of 1 if the answer on the scale is more than 4 (positive influence), 0 if the answer is 4 (no influence), and -1 if the answer is less than 4 (negative influence). Table A1 summarizes each value of *partner_influence* for each treatment. The results show notable differences from the findings in our regression analysis. The perceptions of negative influence from quitters did not appear to be strong. For example, an equal proportion of subjects in the high-quitter condition perceived a positive influence (24.05%) and a negative influence (24.05%) on their number of words studied. On the other hand, the perceptions of positive influence from finisher appeared to be strong. For example, 51.90% of subjects in the similar-finisher condition perceived a positive influence in their number of words studied, compared to only 3.08% who perceived a negative influence.

Study 2

Although the tendency was slightly weaker than in Study 1, Study 2 still exhibited a similar pattern where the perception of the influences did not strongly correlate with the number of studied words.

Table A1

Perception of Partners Influences (Study 1)

	similar-quitter	high-quitter	similar-finisher	high-finisher	all treatments
partner_influence					
Negative	23 (31.08%)	19 (24.05%)	3 (3.80%)	10 (14.08%)	55 (18.15%)
No influence	26 (35.14%)	41 (51.90%)	35 (44.30%)	20 (28.17%)	122 (40.26%)
Positive	25 (33.78%)	19 (24.05%)	41 (51.90%)	41 (57.75%)	126 (41.58%)
Total	74	79	79	71	303

Specifically, we found a positive correlation between the number of studied words and the perception of the influences in the similar-quitter and the similar-finisher treatments ($r = .28$, $p = .01$ in the similar-quitter; $r = .46$, $p < .001$ in the similar-finisher). However, this correlation was not present in the other treatments ($r = .14$, $p = .21$ in the high-finisher; $r = .13$, $p = .20$ in the high-quitter). These findings suggest that the perception of the partners' influences does not reflect the actual perseverance well.

Table A2 summarizes each value of *partner_influence* for each treatment as in Study 1. The overall results are consistent with those of Study 1. The perceptions of negative influence from quitters did not appear to be strong. For example, a similar proportion of subjects in the high-quitter condition perceived a positive influence (34.04%) and a negative influence (30.85%) in their number of words studied. Conversely, the perceptions of positive influence from finisher appeared to be strong. For example, 51.95% of subjects in the high-finisher condition perceived a positive influence in their number of words studied, compared to only 6.49% who perceived a negative influence.

Table A2

Perception of Partners Influences (Study 2)

	similar-quitter	high-quitter	similar-finisher	high-finisher	all treatments
partner_influence					
Negative	26 (33.33%)	29 (30.85%)	5 (6.94%)	5 (6.49%)	65 (20.25%)
No influence	31 (39.74%)	33 (35.11%)	33 (45.83%)	32 (41.56%)	129 (40.19%)
Positive	21 (26.92%)	32 (34.04%)	34 (47.22%)	40 (51.95%)	127 (39.56%)
Total	78	94	72	77	321

Discussion

Our analysis of the perception of partners' influences on perseverance, obtained from the survey responses, suggests that these perceptions do not strongly correlate with actual perseverance. In other words, subjects misjudged their partner's influences on the number of words they studied. Interestingly, the results also imply a *positive perception bias*. Specifically, among subjects paired with low perseverance peers, the majority believed they perceived no influence or even a positive influence from their partners, despite pronounced negative effects actually being observed across both studies. In contrast, among those paired with high perseverance partners, the majority reported positive influences, although non-significant positive effects were actually observed in our studies.

These findings imply that perception measures, commonly used in previous studies (e.g., Keefe, 1994; Steinberg & Silverberg, 1986; Urberg et al., 1990), do not accurately reflect subjects' actual behaviors and highlight the importance of directly measuring behaviors.