The Influence of University Apparel on Stranger Interaction

JULIAN STRIETZEL, Technical University of Munich, Germany
SRISHA RAJ, University of California, Berkeley, United States
VETLE WIEDSWANG JAHR, University of Oslo, Norway
WILLIAM KUCHENBUCH, University of Michigan, Ann Arbor, United States

Additional Key Words and Phrases: Experimentation, Causal Inference, Social Bonding on Campus, Flyering, University Branding, University Identity

ACM Reference Format:
Julian Strietzel, Srisha Raj, Vetle Wiedswang Jahr, and William Kuchenbuch. 2024. The Influence of University Apparel on Stranger Interaction. 1, 1 (December 2024), 19 pages.

1 INTRODUCTION

Social integration in large universities can be challenging, with many students feeling isolated despite being surrounded by peers. Previous research has shown that shared identifiers, such as school pride or common experiences, can foster a sense of belonging [4, 5]. We hypothesize that university-branded merchandise could act as a subtle social cue that promotes interaction between strangers. To test this hypothesis, we conducted a field experiment on the effect of branded merchandise on flyering success. Student flyer distributors were randomly assigned to wear university-branded or neutral non-branded clothing to investigate whether wearing university-branded clothing would have an effect on acceptance of flyers. By comparing the acceptance rate of people passing by under the two conditions of this randomized control trial (RCT), our aim is to determine whether university-branded apparel could work as a low-cost signal of shared identity, thus improving flyer acceptance. Flyer acceptance serves as a proxy measure for gauging how small in-group signals improve social interactions between strangers in a university setting.
Data and Analysis can be found on Drive and GitHub.

Our key contributions are (1) conducting a field experiment based on a randomized control trial, (2) gaining insights into campus dynamics, engagement and social bonding, and (3) making our data and insights available to ensure transparency and foster future work on this topic.

1.1 Research Question

We investigate the following research question:

> ***What effect does university-branded merchandise have on flyering acceptance rates on a college campus?***

Authors' addresses: Julian Strietzel, Technical University of Munich, Munich, Germany, julian.strietzel@tum.de; Srisha Raj, University of California, Berkeley, Berkeley, United States, srisharaj@berkeley.edu; Vetle Wiedswang Jahr, University of Oslo, Oslo, Norway, vetlewj@uio.no; William Kuchenbuch, University of Michigan, Ann Arbor, Michigan, United States, willkuc@umich.edu.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.
© 2024 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

We were additionally interested in the difference of effects between individuals.

Flyering is a traditional method of non-commercial advertising commonly used on university campuses. Flyers are typically distributed in high-traffic areas to students, providing information about events, clubs, or activities, and fostering community interaction.

In our study, the **Flyer Acceptance Rate** is defined as the proportion of students passing by who accept a flyer. We chose to measure social interaction as a proportion to account for variations in the number of students passing by during specific time periods — such as around class start and end.

Additionally, we consider how major campus events (e.g., college football games) influence group affiliation and flyering acceptance rates when wearing rival university-branded merchandise.

1.2 Background and Motivation

Shared group identity has been shown to foster a sense of belonging and increase prosocial behavior, even across demographic differences, such as ethnicity [4]. This experiment investigates whether wearing university-branded apparel can create an in-group identity among strangers, thereby influencing engagement behaviors such as accepting flyers from a stranger.

Clothing is particularly relevant as it serves as a accessible visual signal of group membership and can shape social interactions. College students often use clothing to express their identity and signal belonging [5]. Purchasing university-branded merchandise is closely tied to a student's attachment to their school [7] which is especially dominant on American college campuses. Beyond social identification, the signaling effect of clothing may have external benefits. For instance, studies have shown that individuals dressed neatly received more donations than those dressed messily, highlighting the impact of appearance on behavior [6].

We hypothesize that wearing university-branded clothing while distributing flyers will improve flyer acceptance rates by fostering a sense of shared identity and in-group belonging between the distributor and recipient. This approach provides a mutually beneficial, low-cost strategy to increase campus engagement. Prior research emphasizes the importance of student involvement for positive college outcomes [2, 8].

Furthermore, tangible forms of communication — such as flyers — can leave lasting impressions. A study by Royal Mail and FreshMinds [3], found that 89% of adults remember receiving advertisements through door drops, demonstrating the higher retention effect of physical materials than other forms of communication tested. Although door drops differ from campus flyering, the advantages of a tangible leaflet remain. Therefore, flyering serves as an effective proxy for social bonding, as it maintains high retention even during short interaction periods.

This project investigates how signaling university affiliation through clothing affects the acceptance rate of flyers between strangers on campus, specifically within the unique environment of UC Berkeley. We hope to not only explore the social factors affecting student interaction, but also provide insight into strategies for fostering stronger campus engagement.

1.3 Rationale for UCB Campus Suitability

UC Berkeley's campus is a good fit for this experiment for the following reasons:

(1) Flyering is a common practice among student organizations, especially in high-traffic areas such as Sproul Plaza. These hubs of student activity provide a natural environment to observe social interactions between strangers with a shared identity.

(2) The student body regularly wears university-branded merchandise, making the result of this experiment relevant to students — especially those flyering.
(3) UC Berkeley's long-standing rivalry with Stanford University offers a unique opportunity to study group identity. This is especially evident during campus traditions like the annual "Big Game" football match against Stanford.

2 PILOT EXPERIMENT

Before conducting the main experiment, we carried out a pilot experiment to gather preliminary data to inform our power analysis, address potential challenges, and refine our experimental design. Through the pilot, we identified and addressed the following critical issues before finalizing our methodology:

(1) **Measurement Standardization**: Ensuring consistent methods for counting individuals passing by and accepting flyers.
(2) **Flyering Methods**: Having a pre-defined script and approach when distributing flyers for uniform interactions
(3) **Location Selection**: Selecting narrower flyering locations to improve accuracy in counting of individuals
(4) **Refining Metrics**: Removing QR code scans as a metric, and focusing exclusively on the proportion of flyers accepted

In addition to improving our experimental design, the pilot experiment allowed us to collect preliminary data to be used for our power analysis. By examining these preliminary results, we were able to estimate the number of sessions required and the minimum effect size required to reject the null hypothesis.

2.1 Methodology

Within the *treatment condition* students were wearing bright blue and yellow UC Berkeley overalls, along with branded t-shirts, sweaters, or caps. We avoided wearing standard UC Berkeley merchandise (e.g., hoodies, sweatshirts) in the treatment condition because they are commonly worn by students, and were therefore considered to not provide a strong visual cue. Under the *control condition* experimenters were wearing neutral, unbranded clothing in neutral tones. We chose to flyer for *Paws for Mental Health*, a non-political, broadly supported club that connects students with dogs to support mental well-being.

Each session was 10 minutes, and the goal was to distribute 10 flyers per team member. We worked in pairs for all sessions. Additionally, each set of flyers per session had a unique QR code. Along with measuring the flyer acceptance rate, we included a QR code scan rate to gauge continued interest beyond simply accepting a flyer. We alternated between treatment and control conditions and conducted flyering in high-traffic areas like Sproul Plaza and Sather Gate to ensure randomization and reach a diverse student population. Our collected pilot data can be is attached in the appendix Table 5.

Early in the experiment, we observed that many individuals had already received a flyer from the other team, leading to repeated exposure to the treatment. To minimize this spillover effect and considering pedestrian flow on campus, one team relocated to Memorial Glade after completing a session at the main locations to guarantee space from each other. We wanted to test two different locations to estimate the effect of this covariate. However, switching locations to a smaller, quieter, and more dense area did not completely resolve the problem. Although this quieter area allowed individuals to be more receptive to the flyers, it also made the overalls stand out more; this caused individuals to avoid interaction by changing their path or by visibly looking at their phones.

Furthermore, we observed that wearing distinctive UC Berkeley merch positively influenced our confidence, as it fostered a sense of being legitimate members of the community. As we conducted more sessions, our fear of rejection decreased, and we were more enthusiastic to approach individuals with flyers. We noted that distributing flyers per 10-minute session was completed quickly, with our teams finishing in an average of 5 minutes per session. Based on these findings from the pilot, we refined our analysis by focusing on shorter time spans per session and applying a proportion metric to account for variations in campus traffic flow. These observations from the pilot phase likely influenced our results and led us to consider a more standardized method for future analysis.

2.2 Power Analysis

Prior to conducting our experiment, we performed a power analysis to determine the appropriate sample size. Our assumptions were informed by interviews with four students who regularly distribute flyers on Sproul Plaza (three from the Triathlon club and one from 'Glamour Girls'), along with relevant literature on clothing effects in social interactions. Based on these interviews, we estimate a baseline distribution rate of approximately 5 flyers per 10-minute session (standard deviation = 0.5) and 1 QR code scan per session (standard deviation = 0.25) in the control condition.

Drawing from research on clothing effects in donation recruitment (Levine, Bluni, Hochman 1998), which found a threefold increase in donation rates between different clothing conditions, we conservatively estimated a treatment effect of 2.5 additional flyers per session and 0.5 additional QR code scans when wearing Cal merchandise. We simulated three scenarios: (1) our expected effect sizes, (2) more conservative effect sizes (1 additional flyer, 0.25 additional scans), and (3) our expected effect sizes with higher variance (doubling the standard deviations).

Using Monte Carlo simulations with 300 repetitions per sample size, we calculated statistical power across different sample sizes for both outcome measures. With our expected ATEs and standard deviation (Scenario 1), a sample size of n=12 observations should achieve > 80% power for the flyer handouts, and ~ 80% power for the QR codes. With more conservative ATEs, n=12 should still achieve > 80% for the flyer handouts. The results of the Monte Carlo simulations are shown in Figures 1, 2, 3.

2.3 Analysis

Our primary outcome measures for the pilot experiment included the raw count of flyers successfully distributed and the number of subsequent QR code scans from these distributed flyers. The dataset further includes three key variables: *Treatment*, a binary indicator (1 or 0) denoting whether the distributor wore Cal merchandise; *Session*, marking the specific timestamp of each distribution period; and *Team*, identifying which distributor team handed out the flyers at a specific location.

As our flyers ran out before the ten-minute limit, we recorded the actual time it took to distribute all ten flyers. Using these data, we extrapolated the theoretical number of flyers we could have distributed in the full ten minutes, as well as the corresponding number of QR code scans. We calculated these extrapolated values by dividing the measured quantities by the fraction of the 10-minute period we used. For example:

$$\text{Extrapolated Value} = \frac{\text{Measured Value}}{\frac{\Delta t}{10\text{ minutes}}}$$

We based our analysis on those extrapolated values. This allowed us to estimate the performance as if the activity had continued for the planned duration.

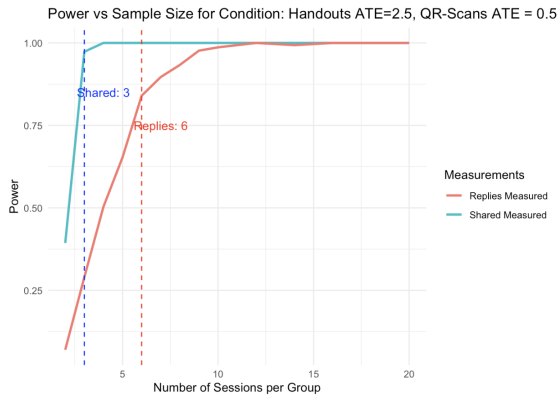
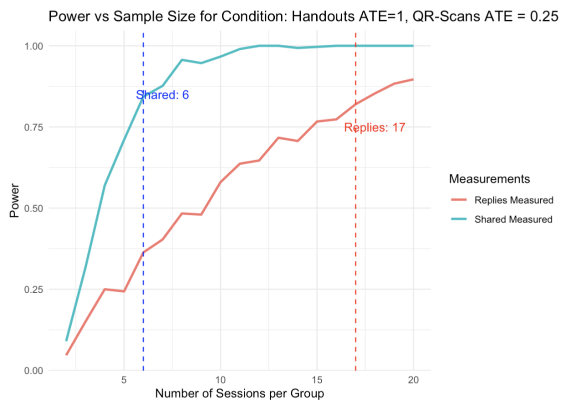
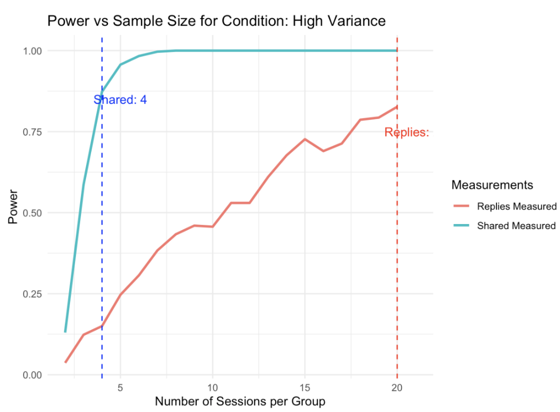


Fig. 1. Power Simulation for Scenario 1.



1

Fig. 2. Power Simulation for Scenario 2.



2

Fig. 3. Power Simulation for Scenario 3.

3

For our analysis, we employed linear regression models to estimate the average treatment effect (ATE). Linear regression was an appropriate choice because it allowed us to easily control for team-specific effects by including Team as a covariate. Additionally, linear regression provides easily interpretable coefficients representing the change in distribution and scan rates attributable to wearing Cal merchandise.

Our initial analysis employed a simple treatment-control comparison through the following specification:

$$Y_i = \beta_0 + \beta_1 \text{Treatment}_i + \varepsilon_i$$

where Y_i represents either the number of flyers distributed or QR codes scanned for observation i, and $Treatment_i$ is a binary indicator for wearing Cal merchandise. This basic model reveals no statistically significant difference between treatment and control conditions for either outcome measure (see table 1).

To improve precision by accounting for potential location and distributor characteristic effects and their interaction with the treatment, we expanded our model to include team fixed effects and their interaction with the treatment:

$$Y_i = \beta_0 + \beta_1 \text{Treatment}_i + \sum_{j=1}^{J} \theta_j \text{Team}_j + \sum_{j=1}^{J} \lambda_j (\text{Treatment}_i \times \text{Team}_j) + \varepsilon_i$$

where $Team_{ij}$ represents indicator variables for each team j operating at a specific location. The results of this model are shown in Table X. Even with these additional controls and interaction terms, we found no significant treatment effect. The interaction terms between treatment and teams were also not statistically significant, suggesting that wearing Cal merchandise did not differentially affect flyer distribution or QR code scan rates across different locations.

Table 1. Number of flyers distributed and QR codes scanned on Treatment - Experiment 1

	Dependent variable:			
	Flyers distributed		QR codes scanned	
	(1)	(2)	(3)	(4)
Treatment	−0.755	0.842	−0.743	0.439
	(3.274)	(5.112)	(1.465)	(2.133)
Team		1.987		2.551
		(5.112)		(2.133)
Treatment * Team		−3.195		−2.364
		(7.229)		(3.016)
Constant	17.021***	16.027***	2.470**	1.195
	(2.315)	(3.614)	(1.036)	(1.508)
Observations	12	12	12	12
R^2	0.005	0.030	0.025	0.174
Adjusted R^2	−0.094	−0.333	−0.072	−0.136
Residual Std. Error	5.671 (df = 10)	6.260 (df = 8)	2.538 (df = 10)	2.612 (df = 8)
F Statistic	0.053 (df = 1; 10)	0.084 (df = 3; 8)	0.257 (df = 1; 10)	0.560 (df = 3; 8)

Note: *p<0.1; **p<0.05; ***p<0.01
Values extrapolated based on the time it took to distribute 10 flyers per session

3 EXPERIMENT 2

Building on the insights gained from our pilot experiment, we conducted a finalized experiment flyering for *Climate Interactive*, an organization focused on simulating climate solutions. While recognizing that environmental issues may not be entirely politically neutral, we believed they hold broad appeal and would yield meaningful results. Additionally, to avoid potential spillover effects from our previous flyering sessions, we selected this organization for our final study.

3.1 Methodology

From our pilot experiment, we implemented the following improvements:

(1) On **Measurement Standardization**, each researcher used a counter through an application called *GP Counter*. This interface allowed us to have two counters - one for all those who passed us and the other for those who passed us, and accepted our flyers.
consistent methods for counting individuals passing by and accepting flyers.
(2) **Flyering Methods**: Having a pre-defined script and approach when distributing flyers for uniform interactions
(3) **Location Selection**: Selecting narrower flyering locations to improve accuracy in counting of individuals
(4) **Refining Metrics**: Removing QR code scans as a metric, and focusing exclusively on the proportion of flyers accepted

3.2 Analysis

Our outcome measure is the success rate of flyer distribution, calculated as the number of flyers successfully handed out divided by the total number of people passing by. We chose this metric over raw handout counts to control for natural variations in foot traffic, particularly during high-volume periods such as class dismissal times at UC Berkeley, thus reducing the variance in our dataset. The dataset includes four key variables: *Treatment*, a binary indicator (1 or 0) denoting whether the distributor wore Cal merchandise; *Session*, an increasing variable denoting which session the data was recorded in; *Person*, an important co-variate factor identifying which researcher distributed the flyers, included to analyze potential heterogeneous treatment effects and control for distributor-specific variations; and *Team*, a co-variate identify which distributor team the flyer was handed out by. By including *Person* as a variable, we can examine whether wearing Cal merchandise affects different researchers' success rates differently, while also controlling for any systematic differences in distribution effectiveness between researchers. The data was generated by splitting and aggregating the raw data based on timestamps (every 2 minutes), according to when we switched from control to treatment condition (see Appendix 6).

To verify the effectiveness of our randomization procedure, we conducted a co-variate balance check between our treatment (wearing Cal merchandise) and distributor assignments. If randomization was successful, we would expect no systematic relationship between who was distributing flyers and whether Cal merchandise was worn during their sessions. We examined this by testing for association between the treatment variable and the distributor indicator (person factor) with a chi-square test. The results showed no significant imbalance (at $p > 0.8$; see Table 2), confirming that our randomization protocol was properly implemented - distributors were equally likely to be assigned to treatment (wearing Cal merchandise) and control conditions.

For our analysis, we employed linear regression models to estimate the average treatment effect (ATE). We chose linear regression as method, as it (1) allowed us to easily control for distributor-specific effects by including Person as a co-variate, (2) our outcome variable (success rate) followed approximately a normal distribution, as demonstrated in

Treatment	Person J	Person S	Person V	Person W	Total
Control (0)	28	12	10	20	70
Treatment (1)	25	8	9	24	66

Note: χ^2 = 1.5898, df = 4, p-value = 0.8106

Table 2. Distribution of Treatment Assignment Across Distributing People for Co-variate-Balance Check

Figure 4, and (3) it provides easily interpretable coefficients representing the change in success rate attributable to our treatment.

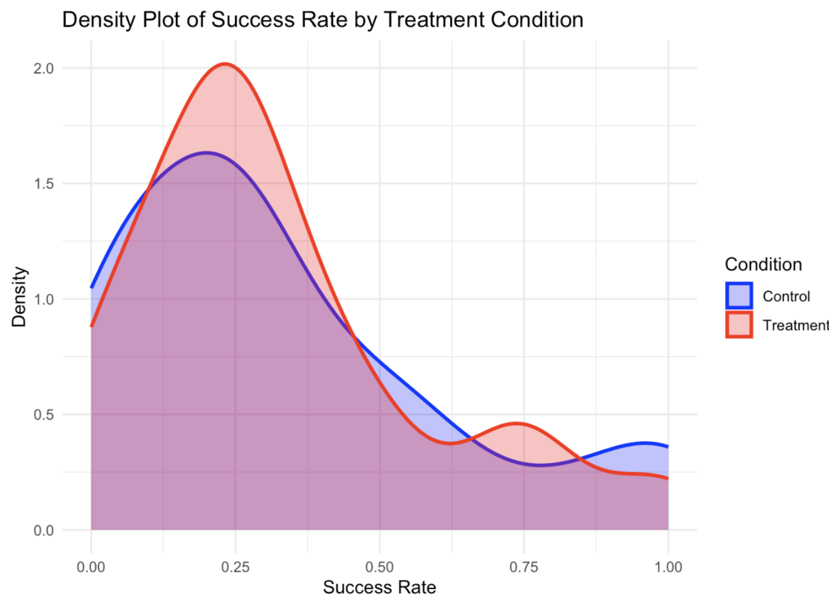


Fig. 4. Distribution of Success Rate By Treatment Condition

Our initial analysis (see Table 3) employed a simple treatment-control comparison, regressing success rate solely on the treatment indicator, through the following specification:

$$Y_i = \beta_0 + \beta_1 \text{Treatment}_i + \epsilon_i$$

where Y_i represents the success rate for observation i, and Treatment$_i$ is a binary indicator for wearing Cal merchandise. This basic model revealed no statistically significant difference between treatment and control conditions. To improve precision by accounting for potential distributor effects and their interaction with the treatment, we expanded our model to include person fixed effects and an interaction term:

$$Y_i = \beta_0 + \beta_1 \text{Treatment}_i + \sum_{j=1}^{J} \gamma_j \text{Person}_{ij} + \sum_{j=1}^{J} \delta_j (\text{Treatment}_i \times \text{Person}_{ij}) + \epsilon_i$$

The results of this model are shown in Table 3. Even with these additional controls and interaction terms, we found no significant treatment effect. The interaction terms between treatment and individual distributors were also not statistically significant, suggesting that wearing Cal merchandise did not differentially affect different distributors' success rates.

To explore potential location-specific effects, we conducted an additional analysis at the team level, where each team consisted of two distributors operating at the same location. This was estimated by the model:

$$Y_i = \beta_0 + \beta_1 \text{Treatment}_i + \sum_{k=1}^{K} \theta_k \text{Team}_{ik} + \sum_{k=1}^{K} \lambda_k (\text{Treatment}_i \times \text{Team}_{ik}) + \epsilon_i$$

where $Team_{ik}$ represents indicator variables for each team k operating at a specific location. This specification allowed us to examine whether the treatment effect varied by location while controlling for team-specific characteristics. Consistent with our previous findings, we found no significant treatment effect, and the interaction terms between treatment and teams were also not statistically significant. This suggests that wearing Cal merchandise did not improve flyer distribution success rates even when accounting for location-specific factors.

To address the statistical power limitations, we conducted an additional analysis using unclustered data, where each potential flyer recipient was treated as an individual observation. In this analysis, success was coded as 1 if the individual accepted a flyer, and 0 if they either declined or passed by. This approach allowed us to disaggregate the data into the 1.5 thousand observations we actually acquired.

The unclustered analysis revealed a significant treatment effect at the $p < 0.1$ level, though it was not significant at the $p < 0.05$ level (see Table 4). Furthermore, two of the interaction terms between treatment and individual distributors were statistically significant, indicating a negative treatment effect for those two distributors, suggesting that wearing Cal merchandise reduced their success rate. However, for the other two distributors, the interaction terms were not significant, showing no deviation from the overall treatment effect, which exhibited a slight positive trend. Thus, while this additional analysis uncovers some insights, it does not provide conclusive evidence of a treatment effect. These findings highlight the importance of further investigation into the treatment's interaction effects, as they vary significantly across individuals and might cancel each other out when aggregated (see Model 1 vs. Model 3 in Table 4), possibly even pointing in opposing directions.

In interpreting these results, several statistical considerations must be taken into account. First, our randomization procedure was not applied individually, but rather on a per-session basis, with sessions lasting ten minutes. This introduces clustering within the data, meaning that we are not perfectly handling randomization at the individual level. Nevertheless, we corrected for this issue in our previous analyses by ensuring that multiple groups, with opposing treatments, were handing out flyers simultaneously across multiple sessions throughout the experiment. Therefore, it could be argued that using only clustered standard errors would lead to inflated standard errors, as suggested by Abadie et al. (2022) [1]. Since the treatment (wearing Cal merchandise) is applied at the individual level, each individual's outcome (whether they accept a flyer) can be considered as a separate observation. If we assume that the treatment effect varies from person to person, then focusing on individual-level data might negate the need for clustering as the treatment is not a group-level intervention. Therefore the estimates can still be considered un-biased.

Additionally, we observed potential spillover effects within sessions. Specifically, our analysis captures both the direct treatment effects and any spillover effects, as each individual's decision was treated independently. Spillover effects in this context would arise if one person's decision to accept or reject a flyer influenced the behavior of nearby pedestrians. It is important to note that these spillovers occur primarily within groups, as we anticipated intra-group spillovers in our design. Although these spillovers can influence our results, they are not an unwanted artifact; rather,

Table 3. Success Rate on Treatment - Experiment 2

	Dependent variable:		
	Success Rate		
	Basic	Interacted Team	Interacted Person
	(1)	(2)	(3)
Treatment	−0.010	−0.050	0.008
	(0.041)	(0.062)	(0.075)
Team vw		−0.004	
		(0.058)	
Treatment*Team vw		0.069	
		(0.082)	
Person j			0.044
			(0.094)
Person s			−0.009
			(0.101)
Person v			-0.156^{*}
			(0.080)
Person w			−0.011
			(0.080)
Treatment*Person j			−0.042
			(0.146)
Treatment*Person s			**−0.124**
			(0.146)
Treatment*Person v			**0.068**
			(0.112)
Treatment*Person w			−0.038
			(0.112)
Constant	0.333^{***}	0.335^{***}	0.365^{***}
	(0.029)	(0.042)	(0.052)
Observations	179	179	179
R^2	0.0003	0.007	0.046
Adjusted R^2	−0.005	−0.010	−0.005
Residual Std. Error	0.273 (df = 177)	0.274 (df = 175)	0.273 (df = 169)
F Statistic	0.057 (df = 1; 177)	0.432 (df = 3; 175)	0.898 (df = 9; 169)
Note:			*p<0.1; **p<0.05; ***p<0.01
			Success rate is the cumulated success rate per virtual 2 minute session
			Treatment constant is for person js, which was partly not marked in the data

they represent a relevant impact that should be accounted for when applying the treatment in real-world settings. These spillovers are included in the clustered analysis, further enriching our understanding of the treatment's effects.

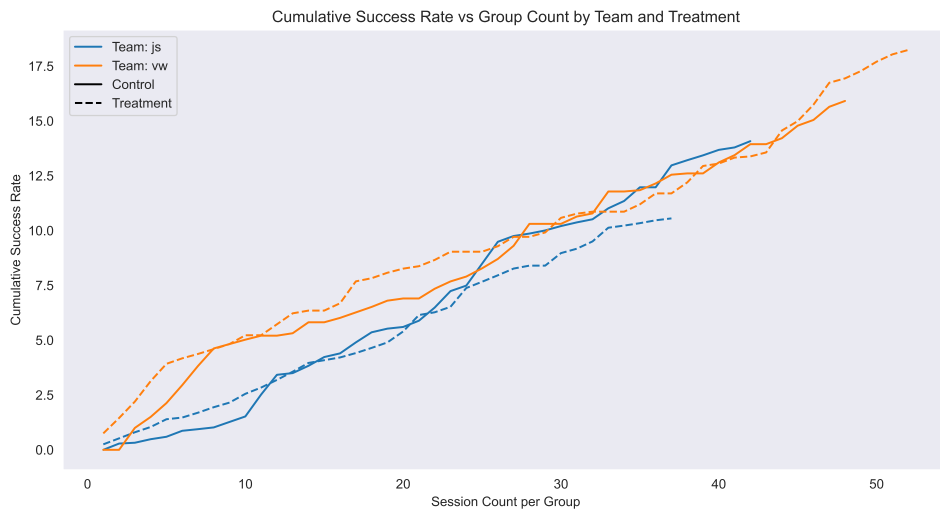
While the individual-level analysis does not definitively establish the treatment effect, it provides valuable insights and suggests directions for further research. It underscores the need for more granular exploration of how treatment effects interact across different individuals and contexts.

Table 4. Success Rate on Treatment (ignoring clustering) - Experiment 2

	Dependent variable:		
	Success Rate		
	Basic	Interacted Team	Interacted Person
	(1)	(2)	(3)
Treatment	0.005	0.028	**0.057***
	(0.021)	(0.028)	(0.032)
Team vw		0.147***	
		(0.030)	
Treatment*Team vw		−0.059	
		(0.042)	
Person j			−0.044
			(0.050)
Person s			−0.061
			(0.050)
Person v			−0.147***
			(0.043)
Person w			0.023
			(0.042)
Treatment*Person j			−0.015
			(0.074)
Treatment*Person s			**−0.131***
			(0.068)
Treatment*Person v			−0.015
			(0.061)
Treatment*Person w			**−0.143****
			(0.059)
Constant	0.288***	0.225***	0.318***
	(0.015)	(0.020)	(0.022)
Observations	1,855	1,855	1,855
R^2	0.00004	0.017	0.022
Adjusted R^2	−0.001	0.016	0.017
Residual Std. Error	0.454 (df = 1853)	0.451 (df = 1851)	0.450 (df = 1845)
F Statistic	0.066 (df = 1; 1853)	10.978*** (df = 3; 1851)	4.565*** (df = 9; 1845)

Note: *p<0.1; **p<0.05; ***p<0.01
Regressing treatment per individual person passing, on them taking the flyer or not
This ignores the clustered randomization of the sessions
Treatment constant is for person js, which was partly not marked in the data

The following is a visual representation of our findings, showing a cumulative total success rate over time, between conditions and distributors. If there were a significant treatment effect, the treatment and control lines would continuously diverge. This graph again indicates that there is no general difference between control and treatment conditions. It emphasizes though, that treatment effects vary between individuals, and might hide some existing underlying effect.



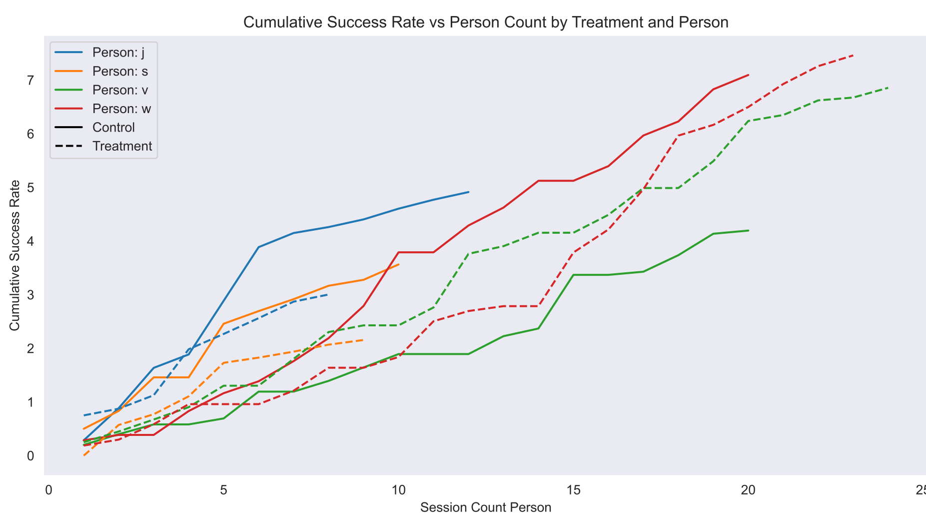


Fig. 5. Cumulative sum of success-rate over sessions per group and person

4 DISCUSSION

In investigating whether university-branded merchandise significantly affects flyer acceptance rates on UC Berkeley's campus, our study found no statistically significant effect at the group level (Table 4). At the individual level, researchers 's' and 'w' demonstrated statistically significant negative interaction effects (-0.131 and -0.143, respectively). Additionally, Figure 5 illustrates that all researchers, except for researcher 'v', exhibited a negative divergence in cumulative success rates under the treatment condition, compared to the control condition. While these individual-level patterns are important to consider, their significance diminishes after properly accounting for the clustering effects of the sessions. Nevertheless, these findings highlight potential underlying dynamics that may not be captured at the group level, especially given the limited number of sessions we conducted.

The randomization of sessions in this study highlights the importance of accounting for contextual factors when evaluating flyer acceptance rates. Our findings suggest that individual researchers may have a greater impact on acceptance rates than the clothing condition, as significant treatment effects were observed at the individual level but not at the group level. To validate this hypothesis, larger sample sizes and additional controlled experiments are required.

Furthermore, these results align with prior work by Osbaldiston and DeBoer, which emphasizes the context-dependent nature of clothing's influence on social behavior [6]. For example, their study found no significant difference in donation amounts between individuals dressed in Santa costumes and those in suits. While clothing can shape social interactions, contextual factors—such as walking in a group or interacting with someone of the same gender—may play a more critical role than clothing alone.

4.1 Limitations

While the findings from this study provide insights into the effect of university-branded clothing on flyer acceptance rates at UC Berkeley, several limitations restrict the broader applicability of these findings.

First, in-group flyering is relatively niche to a college setting, as members often do not need to campaign for causes. Moreover, the effects we observed may reflect unique aspects of UC Berkeley's campus culture—such as the widespread use of school merchandise and the intense Stanford rivalry, as discussed in Section 1.3. The study's limited number of sessions and single location further restrict our ability to draw conclusions about social interactions in different contexts or populations.

Group dynamics presented a significant challenge in this study. Often, when one individual in a group accepted or rejected a flyer, their decision influenced others in the group. In many cases, one person would take a flyer on behalf of the rest, resulting in shared engagement rather than individual responses. This dynamic complicated our assumption that treatment effects occur at the individual level. Since these clusters were neither predefined nor consistent, our analysis could not fully account for group-level effects. We also encountered spillover effects, where some participants passed by one flyering group before encountering the other. This exposure may have biased their responses, despite our efforts to mitigate the issue in the pilot experiment by increasing the distance between groups. While these adjustments helped, they did not completely resolve the problem.

To strengthen the treatment effect, we chose to wear distinctive UC Berkeley overalls that are bright blue and yellow striped, as seen in Figure 6. While these overalls heightened the visibility of the treatment, they differ from the university-branded apparel typically worn by students, such as t-shirts or sweaters. This distinction makes it difficult to generalize our findings to more conventional forms of university-branded clothing.

The study also faced logistical limitations. We were unable to test the treatment condition involving rival university-branded clothing (Stanford) during the 'Big Game' due to heavy rain. Although we attempted to flyer in heavy rain, the campus was empty and the flyers were immediately damaged when exposed to the rain. Thus, we had to cancel this portion of the experiment. We had hypothesized that students would feel a stronger identity and in-group belonging to their home university in rival settings like this, hence increasing the treatment effect for both wearing UC Berkeley and Stanford clothing.

Factors such as time of day—particularly in relation to proximity to class—and gender seemed to influence both the proportion of students who walked past and those who accepted the flyers. However, we did not block for these variables, which could have helped reveal any heterogenous treatment effects within UC Berkeley's diverse student body and provided a more nuanced understanding of the treatment condition. Unfortunately, implementing such blocking would have required speculative assumptions, as any attempt to collect information would have needed to occur post-treatment. This method could introduce bias, as students who chose to provide their information likely differ from those who did not.

Fig. 6. Treatment 1: Distributors wearing UC Berkeley-branded apparel

4.2 Future Work

Although our results did not reveal significant differences at the group level, the significant negative treatment effects observed at the individual level open avenues for further exploration. Future research could increase the statistical power of this experiment and enhance the representativeness of the campus sample by conducting additional sessions with a randomized approach. Moreover, since we were unable to flyer during the "Big Game," future studies could investigate the treatment effect of wearing rival university-branded merchandise both in regular settings and during campus tradition events to capture any contextual differences in flyer acceptance rates.

Expanding the diversity of individuals conducting the flyering, particularly by accounting for factors such as gender, could provide valuable insights into how the flyerer's identity influences treatment effects. This would allow for blocking to occur based on characteristics of the researchers rather than solely relying on speculative or post-treatment information about individuals on campus. By controlling for researcher-specific factors, future studies could isolate their potential impact on flyer acceptance rates, enhancing the validity of the treatment effects attributed to the university-branded clothing condition.

Finally, extending the experiment to rival schools—such as flyering at both Stanford and Berkeley—could help determine whether student populations at these institutions respond differently or if the observed clothing condition effect is consistent across campuses. This comparative approach could provide a deeper understanding of the interplay between school identity and social behavior.

5 CONCLUSION

Our research aimed to explore the effect of university-branded merchandise on flyering acceptance rates on a college campus. We hypothesized that a greater sense of in-group belonging and shared identity, signaled by wearing university-branded apparel, would make it easier for strangers to interact with each other at a college campus.

Contrary to our expectations, we observed no significant evidence at the group level indicating that wearing university-branded apparel has a positive effect on flyer acceptance rate. There was however strong and statistically significant evidence for two of the individuals when investigating the results at the individual-level that wearing university-branded apparel had a negative effect on the acceptance rate. This emphasizes that individual differences and other unmeasured factors are important to consider when analyzing the data from our experiments.

An essential part of the process was the iterative approach of refining our experimental design after conducting the pilot experiment. Through our pilot experiment, we learned valuable lessons on standardizing as much as possible so that everything is performed and measured consistently across the distributors. We were also able to use the results from our pilot experiment to perform a power analysis and hence estimate the required sample sizes and effect sizes needed to reject the null hypothesis.

Ultimately, our research highlights that although apparel may serve as a signal of in-group belonging and social identity, how clothing shapes everyday encounters is a more nuanced problem, requiring further research and careful experimental design.

ACKNOWLEDGMENTS

We would like to thank Professor Alex Hughes for his guidance and support throughout this project, particularly in developing our experimental design.

We also acknowledge Paws for Mental Health and Climate Interactive for permitting us to flyer on their behalf, which significantly contributed to the execution of our research methodology.

A DATA

Data can be found on our Google Drive and on GitHub.

Table 5. Data Appendix for Pilot

treat	#flyers	#qr_scanned	sec_to_cap	group	time	date	day_of_week	quote_used	#qr_extra	#flyers_extra
1	10	1	280	0	12:20	10/17	Thurs	0.466667	2.142857	21.428571
0	10	1	283	0	12:35	10/17	Thurs	0.471667	2.120141	21.201413
1	10	0	390	0	12:50	10/17	Thurs	0.650000	0.000000	15.384615
0	10	1	410	0	13:02	10/17	Thurs	0.683333	1.463415	14.634146
1	10	2	435	0	13:15	10/17	Thurs	0.725000	2.758621	13.793103
0	10	0	490	0	13:30	10/17	Thurs	0.816667	0.000000	12.244898
0	10	0	472	1	12:20	10/17	Thurs	0.786667	0.000000	12.711864
1	10	2	382	1	12:35	10/17	Thurs	0.636667	3.141361	15.706806
0	10	3	202	1	12:50	10/17	Thurs	0.336667	8.910891	29.702970
1	10	0	305	1	13:02	10/17	Thurs	0.508333	0.000000	19.672131
0	10	2	516	1	13:15	10/17	Thurs	0.860000	2.325581	11.627907
1	10	2	517	1	13:30	10/17	Thurs	0.861667	2.321083	11.605416

Table 6. Description of Variables in the Experiment Dataset per Session

Variable Name	**Data Format**	**Description**
session_starts	Datetime	Timestamp indicating when the session started (relative to exp start time).
treatment	Binary (0/1)	Indicates whether the record belongs to the treatment group.
person	Categorical	Identifier for the individual conducting the session.
team	Categorical	Identifier for the team involved in the session.
handed_out	Numeric	Number of items handed out during the session.
passed_by	Numeric	Total number of people who passed by during the session (including the people flyered to).
session_within_session	Integer	Nested session identifier for two min subsets within a ten min session.
success_rate	Percentage	Ratio of success to total attempts (0-1).
day	Categorical (Text)	Day of the week (e.g., Monday, Tuesday).
person	Categorical	Identifier for the individual conducting the session.
session	Categorical	Overall session identifier.

Table 7. Description of Variables in the Experiment Dataset for Individuals

Variable Name	Data Format	Description
count	Integer	Count of records for each event/observation.
date	Date	Date of the session (in the format: YYYY-MM-DD).
time	String (Character)	Time of the session in HH:MM:SS.sss format.
team	Categorical	Identifier for the team conducting the session.
timestamp	Datetime ()	Timestamp indicating the exact time when the record was logged.
seconds_after_start	Numeric (Double)	Number of seconds elapsed after the session started.
success	Logical (TRUE/FALSE)	Boolean indicating if the session was successful (TRUE) or not (FALSE).
person	Categorical	Identifier for the individual conducting the session.
treatment	Numeric (0/1)	Binary indicator whether the record belongs to the treatment group (1) or not (0).

B FLYERS

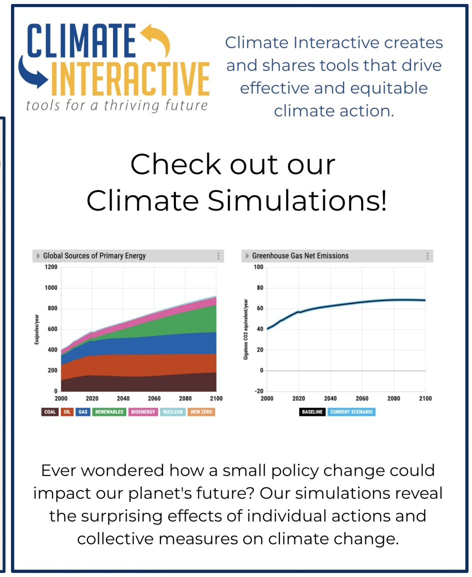


Fig. 7. Used flyers: Paws for Mental Health (Left), Climate Interactive (Right)

REFERENCES

[1] Alberto Abadie, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2022. When Should You Adjust Standard Errors for Clustering? arXiv:1710.02926 [math.ST] https://arxiv.org/abs/1710.02926

[2] Alexander Astin. 1984. Student Involvement: A Development Theory for Higher Education. *Journal of College Student Development* 40 (01 1984), 518–529. https://www.researchgate.net/publication/220017441_Student_Involvement_A_Development_Theory_for_Higher_Education

[3] FreshMinds and Royal Mail. 2011. Challenging pre-conceptions about door drops. https://www.letterboxconsultancy.com/wp-content/uploads/2011/10/Challenging_pre-conceptions_about_door_drops_8_07_11_revised-1-1.pdf

[4] Shana Levin, Stacey Sinclair, Jim Sidanius, and Colette Van Laar. 2009. Ethnic and University Identities across the College Years: A Common In-Group Identity Perspective. *Journal of Social Issues* 65, 2 (2009), 287–306. https://doi.org/10.1111/j.1540-4560.2009.01601.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-4560.2009.01601.x.

[5] Mijeong Noh, Meng Li, Kaleb Martin, and Joseph Purpura. 2015. College men's fashion: clothing preference, identity, and avoidance | Fashion and Textiles | Full Text. https://fashionandtextiles.springeropen.com/articles/10.1186/s40691-015-0052-7

[6] Richard Osbaldiston and Brittany de Boer. 2011. The Effects of Wearing a Costume on Charitable Donations. *Psychological Reports* 108, 1 (Feb. 2011), 167–168. https://doi.org/10.2466/01.07.PR0.108.1.167-168 Publisher: SAGE Publications Inc.

[7] Bomi Park. 2015. What Makes Bearcats Buy? School Bonds, Attitudes toward School, and PI of School Merchandises. https://www.mckendree.edu/academics/scholars/park-issue-25.pdf

[8] Karen L. Webber, Rebecca Bauer Krylow, and Qin Zhang. 2013. Does Involvement Really Matter? Indicators of College Student Success and Satisfaction. *Journal of College Student Development* 54, 6 (2013), 591–611. https://muse.jhu.edu/pub/1/article/528367 Publisher: Johns Hopkins University Press.