

Targeting Nearby Influencers: The Acceleration of Natural Triadic Closure by Leveraging Interconnectors

Journal of Marketing

1-20

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00222429231223420

journals.sagepub.com/home/jmx



Jacob Goldenberg, Andreas Lanz , Daniel Shapira, and Florian Stahl

Abstract

On user-generated content platforms, individuals and firms alike seek to build and expand their follower base to eventually increase the reach of the content they upload. The bulk of the seeding literature in marketing suggests targeting users with a large follower base, that is, high-status influencers. In contrast, some recent studies find targeting lower-status influencers to be a more effective seeding policy. This multimethod article shifts the focus from the follower base of the seeding target to the focal content creator. The authors propose accelerating natural triadic closure by leveraging first-degree followers as interconnectors to target second-degree followers, that is, the *nearby* (low-status) influencers (who are interconnected with the focal content creator). Empirical studies document that this seeding target is much more effective for building and expanding the follower base, compared with targeting influencers who are not interconnected with the focal content creator—that is, the *remote* (both high- and low-status) influencers—by 2,300% and 46%, respectively. These studies on the acceleration of natural triadic closure are augmented by a preregistered field experiment to obtain convergent validity of the findings.

Keywords

influencer marketing, seeding, social networking, unpaid endorsements, user-generated content, triadic closure, data-based simulation, field experiment

Online supplement: <https://doi.org/10.1177/00222429231223420>

In recent years, user-generated content platforms such as SoundCloud, X (formerly known as Twitter), Instagram, and LinkedIn have become omnipresent. On these platforms, individuals and firms alike seek to build and expand their follower base to eventually increase the reach of the content they upload, which can be divided into paid and unpaid endorsements (Goldenberg et al. 2022). We consider the latter, also known as the follow-for-follow approach, which is a popular way to build a follower base, according to Hootsuite (2022), a major platform management tool: “When you follow a user on Instagram, there’s a good chance they will check out your feed and consider following you back.” SoundCloud (2023), for example, explicitly mentions this approach in its help center for content creators: “The best way to gain meaningful followers is to be one yourself.” The academic literature in this domain suggests targeting users with a large follower base, that is, the high-status influencers (e.g., Goldenberg et al. 2009; Hinz et al. 2011; Hughes, Swaminathan, and Brooks 2019), with some exceptions suggesting the targeting of lower-status ones instead (e.g., Lanz et al. 2019).

All previous work has focused on the *influencer’s* follower base (i.e., low vs. high status). However, there is scant attention,

if any, to the follower base of the *individual or firm* (i.e., the content creator). This could be a missed opportunity due to the tendency of the follow-for-follow approach to attract somewhat remote followers who do not necessarily fit in well with the creator’s follower base and are therefore not very engaged. In this article, we shift the focus to the follower base of an individual or firm, that is, the content creator, and explore how such a creator may capitalize on natural network formation in the immediate vicinity to support seeding efforts. We maintain that the creator’s own follower base is an

Jacob Goldenberg is Full Professor of Marketing, Arison School of Business, Reichman University, Israel, and Visiting Professor, Columbia Business School, Columbia University, USA (email: jgoldenberg@idc.ac.il). Andreas Lanz is Assistant Professor of Digital Marketing Analytics, Faculty of Business and Economics, University of Basel, Switzerland (email: andreas.lanz@unibas.ch). Daniel Shapira is Senior Lecturer of Marketing, Department of Business Administration, Ben-Gurion University of the Negev, Israel, and Permanent Adjunct Research Faculty, Mannheim Business School, University of Mannheim, Germany (email: shapirad@bgu.ac.il). Florian Stahl is Full Professor of Marketing, Mannheim Business School, University of Mannheim, Germany (email: florian.stahl@uni-mannheim.de).

overlooked area in the network for finding surprisingly valuable influencers, previously ignored by the literature. In contrast, in the secondary data we use, creators already capitalize on natural network formation as they target 11.24% of their efforts to the immediate vicinity. This will serve as a real-life benchmark for our policy suggestion.

One of the two fundamental processes of natural network formation is triadic closure—a well-established social dynamic in the networks literature (e.g., Van den Bulte and Wuyts 2007)—which is the tendency to become friends if one has mutual friends.¹ In this article, we investigate how unknown, *emerging* content creators can accelerate this fundamental process of triadic closure. Consider a focal creator A as well as users B and C (see Figure 1). Since C follows B, and B follows the focal creator A, the first-degree follower B is an “interconnector” between A (the creator) and the second-degree follower C (the influencer). Then, if A targets C with an outbound activity (e.g., a follow, comment, like, or private message), and C subsequently follows A back, it results in triadic closure—and this fundamental process was accelerated due to the outbound activity. Along these lines, given an outbound activity from an unknown content creator A, the a priori probability that the *nearby* influencer C (who is interconnected with A through B) will follow A back should be substantially higher than that of a *remote* influencer (who is not interconnected with A). This is our focus.

Since the seeding literature in marketing traditionally focuses on the follower base of the influencer (i.e., low vs. high status), this article adds a new, orthogonal dimension, which is the second-degree followership of the influencer to the content creator (i.e., nearby vs. remote). This second-degree followership is identified by the presence of an interconnector (i.e., a first-degree follower), which results in a 2 (nearby vs. remote) \times 2 (low vs. high status) setting with four seeding targets: (1) nearby low-status influencers (where an interconnector is present), (2) remote low-status influencers (e.g., Lanz et al. 2019), (3) nearby high-status influencers (where an interconnector is present), and (4) remote high-status influencers. We argue and show that the dominant seeding policy is the first.

We use secondary data from a worldwide leading audio platform to study unknown creators who seek to increase the reach of the content they upload—songs and podcasts—by expanding their follower base through directing outbound activities such as follows, comments, likes, or private messages to other users of the platform. In this secondary data we find that remote low-status influencers feature a higher a priori follow-back probability, that is, 6.68%, compared with .02% when targeting remote high-status influencers. In line with our expectation, leveraging interconnectors to close triangles is more effective: we find that the a priori follow-back probability of nearby low-status influencers—the ones who are second-degree followers—more than doubles, increasing from 6.68% to 14.37% (for nearby

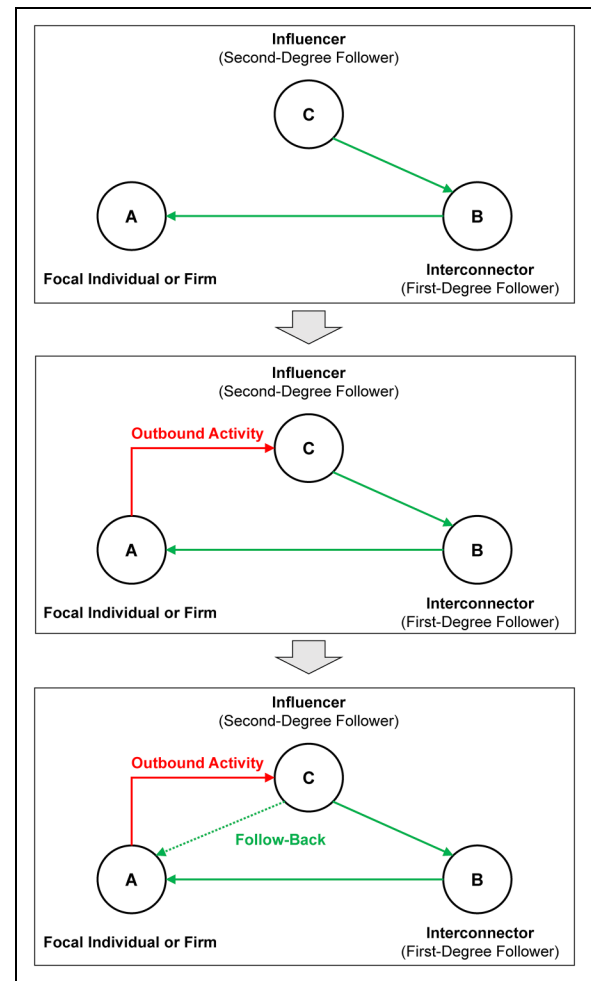


Figure 1. Explanatory Graph.

high-status influencers it increases from .02% to .19% and is thus still extremely marginal). This finding holds when applying propensity-score matching to investigate whether the mere presence of an interconnector indeed makes all the difference, above and beyond the a priori matching between an unknown creator and a low-status influencer (i.e., due to homophily, which is the other fundamental process of natural network formation). We conduct a logistic regression, including various controls and moderators, and find further support for the acceleration of the natural triadic-closure phenomenon.

We refer to this follow-back as a *direct return* since it comes from the influencer directly. In line with Lanz et al. (2019), we also consider the *indirect return*, namely follow-backs coming from the influencer’s follower base as a result of a repost by the influencer. Moreover, there are long-term, *additional indirect returns*: if an influencer is acquired and now follows the unknown content creator, then we also consider cascades into the second- and third-degree follower base due to future reposts by this newly acquired (first-degree) follower and that follower’s followers.

To assess the extent to which the natural triadic-closure phenomenon can support seeding efforts, considering the 2 \times 2 setting, we conduct data-based simulations emulating these

¹ “The friend of my friend is not necessarily my friend, but is far more likely to be my friend than some randomly chosen member of the population” (Newman 2010, p. 199).

rich network dynamics and long-term impacts of the direct, indirect, and additional indirect returns. First, we find that exclusively targeting nearby low-status influencers outperforms targeting remote (and nearby) high-status influencers—the traditional influencer marketing approach—by 2,300% within two years' time. Second, our suggested seeding policy also outperforms targeting remote low-status influencers, in this case by 46%. Third, against a real-life benchmark, it outperforms the unknown creators' actual seeding policy in our data, or status quo targeting, by more than twice. Therefore, we conclude that in the 2×2 setting, low status dominates high status, and nearby dominates remote, such that *targeting nearby low-status influencers is the dominant seeding policy* and is also superior to the status quo targeting. However, the targeting observed in the secondary data by definition involves self-selection. Therefore, to obtain convergent validity of our findings in a more controlled environment, we augment this study with a preregistered field experiment that provides further support for the effectiveness of targeting nearby low-status influencers.

In terms of theoretical implications, we provide (1) a new angle on triadic closure, focusing on the interpersonal level and a pure online setting in the context of seeding, a rarely studied instance due to obvious data-acquisition challenges. Besides trust, which previous work has pointed out, we uncover two additional underlying drivers behind the triadic-closure phenomenon for this setting and context: perceived similarity (i.e., the interconnector creates indirect connectedness between the creator and the influencer) and (favor-)reciprocity (i.e., the interconnector creates a perceived sense of obligation). Furthermore, while adding a new orthogonal dimension to seeding (i.e., nearby vs. remote), namely the mere presence of an interconnector, we provide (2) further empirical support for the effectiveness of targeting low-status influencers, a policy that the seeding literature in marketing has not yet fully recognized (currently only a few works address it; e.g., Ameri, Honka, and Xie 2023; Beichert et al. 2023; Lanz et al. 2019).

In terms of managerial implications, we find that the most straightforward means of acceleration is to simply follow second-degree followers, because everyone can be followed. This expansion of the follower base to increase the reach of the uploaded content can be directly supported by user-generated content platforms, which will have to shift from focusing only on content discovery to offering another recommender system purely for the discovery of seeding targets. At the core of this algorithm should be the leveraging of interconnectors.

Our findings from unpaid endorsements (i.e., based on the follow-for-follow approach) may also have managerial implications for paid endorsements: firms should not underestimate the value of influencers who are (inter)connected with them, and it could be expected that such nearby influencers are associated not only with a higher contract-acceptance rate (Hofstetter, Lanz, and Sahni 2023) but also with higher revenues (Beichert et al. 2023), due to triadic closure (and homophily; Haenlein and Libai 2013). As the seeding literature in marketing has overlooked the firm's own follower base for influencer selection, this article calls for a paradigm shift in this respect.

The remainder of the article is organized as follows. In the subsequent section we discuss the related literature before empirically investigating the effectiveness of our suggested policy against other approaches. We explore model-free evidence and conduct propensity-score matching, a logistic regression accounting for nonlinearities in the estimation, an online experiment, and data-based simulations. Then, we replicate this comparison in the field with an experiment. Finally, we provide both theoretical and managerial implications, discuss our findings, and suggest directions for future research.

Background

This article is related to two streams of literature: (1) influencer marketing and (2) triadic closure. These are, in fact, the two sides of the same coin with regard to the acceleration of natural triadic closure to support seeding efforts.

Influencer Marketing

The motivation for influencer marketing is rooted in the fact that word of mouth affects consumption decisions; hence individuals and firms seek to build and expand their follower base on user-generated content platforms. To achieve this primary goal, they have two options in terms of seeding programs (Haenlein and Libai 2017): paid or unpaid influencer endorsements (Goldenberg et al. 2022).

Paid influencer endorsements refer to the instance in which the firm—or, in rare cases, the individual—pays or incentivizes the influencer to take action (for work on the inherent principal-agent problem, see Hofstetter, Lanz, and Sahni 2023; Pei and Mayzlin 2022; Peng and Van den Bulte 2024). Paid influencer endorsements are usually subject to portfolio considerations across multiple influencers (e.g., Gu, Zhang, and Kannan 2023; Lanz et al. 2023), whereby the specific form of action is defined in the influencer briefing and typically includes (re) posting content, with a variety of factors determining its effectiveness (e.g., Leung et al. 2022; Wies, Bleier, and Edeling 2023).

Unpaid influencer endorsements are the basic form of influencer marketing and refer to the instance in which the individual or firm does not offer compensation but utilizes outbound activities—such as follows, comments, likes, and private messages—to get the influencer's attention (also known as the follow-for-follow approach). The individual or firm can direct such outbound activities to influencers on a user-generated content platform in order to catch their attention, because outbound activities trigger a notification (e.g., Weiler et al. 2022). This basic form of influencer marketing is of high interest, particularly for individuals and small and medium-sized businesses (which make up more than 99% of all businesses in Europe and the United States). They typically lack the monetary budget to pay or incentivize an influencer—since costs are generally high—and thus rely on unpaid endorsements, especially since the seeding objective aligns with that of paid influencer endorsements: to trigger, on the one hand, a consumption

return (i.e., engagement with the content) and, on the other hand, follow-back returns in the form of (1) a direct return (i.e., a follow-back from the influencer) as well as (2) an indirect return (i.e., a repost from the influencer that generates further follow-backs). With unpaid influencer endorsements, the returns are much more uncertain, because there is competition for the influencer's attention (e.g., Gelper, Van der Lans, and Van Bruggen 2021; Rossi and Rubera 2021).

In this article, we consider unpaid influencer endorsements and focus on follow-back returns, because we study unknown creators who seek, by expanding their follower base, to eventually increase the reach of the content they upload, similar to the setting of Ansari et al. (2018). We maintain that unpaid influencer endorsements on user-generated content platforms likely constitute the only feasible option for unknown creators to become known, and this option is widely relied on since the vast majority of creators are unknown (following the typical power-law follower-base distribution prevalent on such platforms). It should be noted that the objective function of known creators, who are outside the scope of this article, can be expected to change with increasing network status: unlike unknown creators, the objective function of known creators may not predominantly be to expand their follower base; rather they focus on keeping the links to the follower base active.

Considering unpaid influencer endorsements, the bulk of the seeding literature in marketing suggests targeting users with a high network status. Note that network status is commonly operationalized by the user's follower base, or indegree (e.g., Wasserman and Faust 1994). Other network information such as the user's betweenness, clustering, or overlap is, however, often very limited or not readily available in most cases and thus is impossible to assess. There are many related terminologies that describe central users (e.g., opinion leaders, market mavens, and influentials), all differing in one or more dimensions. In this article, we adopt the common definition and refer to the status in terms of indegree, following, among others, Hinz et al. (2011), who also adopt this definition: comparing low-status, high-status, and high-betweenness seeding in three studies, they conclude that targeting high-status influencers performs best, a conclusion that also emerges in other studies (e.g., Goldenberg et al. 2009; Hughes, Swaminathan, and Brooks 2019).

In contrast, the seeding literature in marketing has not yet fully recognized the potential of low-status influencers. Following publications that note the value of such seeding targets based on simulations (Watts and Dodds 2007) and analytical work (Galeotti and Goyal 2009), Lanz et al. (2019) present empirical evidence in the context of unpaid influencer endorsements. They show that in many cases targeting low-status influencers is associated with several orders of magnitude higher responsiveness and follow-back probability compared with targeting high-status influencers; hence the associated seeding policy is superior. Beichert et al. (2023) find a similar phenomenon in the context of paid endorsements.

From these works it becomes apparent that the focus is on the influencer's follower base, predominantly low versus high

status. In this article, we add a new orthogonal dimension to seeding by shifting the focus to the follower base of the focal unknown creator (an individual or firm). We argue that there is an overlooked area in the immediate vicinity of the creator populated with valuable influencers.

Triadic Closure

Generally, two processes govern natural network formation: triadic closure and homophily (i.e., intrinsic and contextual factors; Easley and Kleinberg 2010). In this article, we build on the former. In social networks, triangles close due to the natural tendency to become friends if one has mutual friends. This process is distinct from homophily, which revolves around network formation based on similarities. Triadic closure theory goes back to Rapoport (1953), who considers how people will become acquainted in the future, given those they know in the present. Naturally, a common friend of two people in the present implies a likelihood that the two will become friends in the future. Therefore, unlike homophily, triadic closure takes place beyond the individuals' features, or contextual factors, when they eventually form friendship ties. Note that the term "friendship" implies undirected relations. For directed relations, such as "following" on SoundCloud or X, triadic closure is referred to as link copying (Romero and Kleinberg 2010).

One of the important driving forces that underlies triadic closure is trust (Easley and Kleinberg 2010). Triadic closure is a well-established social dynamic, found to be at play on the interorganizational and the interpersonal level such as when individuals are searching for new jobs (e.g., Granovetter 1973; Rajkumar et al. 2022) or when they look for needed products (and receive referrals from their friends; e.g., Schmitt, Skiera, and Van den Bulte 2011; Van den Bulte et al. 2018).

Although very different from the notion of seeding, customer referrals are related to influencer endorsements, because an influencer's (re)post is actually a referral. Therefore, the objective of customer referral programs, which is to target the friends of customers and close triangles, is similar to that of influencer marketing, and our setting in particular. The main difference lies in the key feature that the interconnectors—that is, the first-degree followers—have a passive role. This stands in contrast to referral marketing, where customers are incentivized to actively recommend the firm to their friends. In our setting, the only one who takes action is the focal unknown creator (an individual or firm), by directing outbound activities to other users of the platform to get follow-back returns.

Along these lines, we combine influencer marketing with triadic closure and suggest accelerating this fundamental process by leveraging interconnectors: instead of targeting remote influencers (who are not interconnected with the focal creator), unknown creators should focus on their immediate vicinity and target nearby influencers (who are interconnected), which increases the follow-back probability, especially from the low-status influencers (Lanz et al. 2019). For an overview of the main constructs, definitions, and operationalizations, see Table 1.

Table 1. Constructs, Definitions, and Operationalizations.

Construct	Definition	Operationalization
Creator	A platform user—individual or firm—that seeks to build and expand the follower base to eventually increase the reach of the uploaded content	User who has uploaded at least one content piece (i.e., an audio file)
Low-status influencer	A platform user with a small follower base	User with 0 to 100 followers (i.e., in the lower tail of the follower-base distribution)
High-status influencer	A platform user with a large follower base	User with more than 10,000 followers (i.e., in the upper tail of the follower-base distribution)
Interconnector	A platform user who is a first-degree follower of the creator	User who is a follower of the creator
Nearby influencer	A platform user who is a second-degree follower of the creator	User who is a follower of a follower of the creator
Remote influencer	A platform user who is neither a first- nor a second-degree follower of the creator	User who is neither a follower nor a follower of a follower of the creator
Outbound activity	A follow, comment, like, or private message that is directed from the creator to the influencer and is intended to trigger follow-back returns	Binary; sent or not
Direct return	A follow-back from the influencer directly (meaning the influencer is newly acquired)	Binary; followed back or not
Indirect return	Follow-backs coming indirectly from the influencer's follower base as a result of a repost by the influencer	Count; sum of follow-backs from the follower base of the influencer
Additional indirect return	Follow-backs from cascades into the second- and third-degree follower base due to future reposts by a newly acquired (first-degree) follower and that follower's followers	Count; sum of follow-backs from the cascade

It is important to clarify that this new dimension (nearby vs. remote) is completely different from others relying on Feld's (1991) friendship paradox, which has been utilized to identify and target remote high-status influencers, or at least individuals with higher status (e.g., Kumar and Sudhir 2021; Stephen and Lehmann 2016), and to improve stochastic seeding (e.g., Chin, Eckles, and Ugander 2022). The basic idea behind utilizing the friendship paradox for seeding is to select a random—and thus remote—user and go into the first-degree follower base of this random user to find more connected users (they are on average more connected than the selected random user). In this way, one quickly finds remote high-status influencers. Our proposed approach is the polar opposite: Instead of searching for remote high-status influencers, we examine nearby low-status ones. In terms of implementation, our suggested seeding policy is straightforward. The creator can uncover the first-degree followers by clicking on “followers” in the own user profile, and following the same approach, the creator can then uncover the second-degree followers to find nearby low-status influencers. Therefore, the influencer's key feature is their second-degree followership, based on which the influencer can be identified as indirectly connected, whereas other network measures of closeness (e.g., betweenness) are impossible for creators to assess.

Hypotheses

Consistent with previous literature, triadic closure is a natural process in our secondary data. There is a baseline of followers that unknown creators (with less than or equal to 100 followers) acquire naturally and without targeting, but this baseline is very low compared with what can be achieved when creators direct outbound activities to other users of the platform, namely when targeting low-status

influencers, especially the nearby ones, thereby accelerating natural triadic closure (see also the results of the data-based simulations). We find that 48% of this baseline originates in the second-degree follower base, and 52% comes from all over the platform. Therefore, there is strong evidence that triadic closure, among other processes, governs natural network formation in our data. From this starting point, and building on the preceding theoretical foundation, we expect that unknown creators can accelerate natural triadic closure by targeting nearby low-status influencers.

Based on social networks literature, a second process of natural network formation may play into it, namely homophily, that is, the tendency for follow-backs to occur due to the similarity between the creator and the influencer (Van den Bulte and Wuyts 2007). The core role in triadic closure—although passive—is played by the interconnectors, who serve as silent information transmitters of indirect connectedness between the creator and the influencer. Due to this interconnection, the influencer becomes aware of the second-degree followership with the creator in the network. Moreover, it may increase the perceived similarity between the creator and the influencer, making homophily a contributing factor in the acceleration of the natural triadic-closure phenomenon. Hence, given an outbound activity from the creator to the influencer, the mere presence of an interconnector should increase the follow-back probability, compared with the instance when the interconnector is absent. Naturally, the more interconnectors that are present, the stronger the perception of second-degree followership, which then should even further increase the follow-back probability. Along these lines, we formulate the following hypotheses:

H_{1a}: We expect that the mere presence of an interconnector increases the follow-back probability; that is, this increase is due to triadic closure, above and beyond homophily.

H_{1b}: We also expect that the more interconnectors present, the higher the follow-back probability.

But what role does the (network) status of the influencer play in the interaction with the mere presence of an interconnector? Concerning the follow-back probability, Lanz et al. (2019) find that for an unknown creator, targeting higher-status influencers is associated with a decreasing follow-back probability; hence unknown creators should target low-status influencers. They argue that social identity theory (Tajfel and Turner 1979) is behind this phenomenon; namely, high-status influencers, just like other high-status individuals, exhibit self-focused and self-serving behavior, and this can be attributed to stronger in-group identification, favoritism, and the preservation of group boundaries. Naturally, high-status influencers are constantly targeted with outbound activities; hence there is also competition for attention (Gelper, Van der Lans, and Van Bruggen 2021). Based on this social psychology literature on status and the resulting inter- and intragroup behavior, it follows that the mere presence of an interconnector by definition puts the creator into the same (network) group with the influencer, through the indirect connectedness. In addition, the mere presence of an interconnector helps the creator stand out in the competition for attention. Therefore, we expect that the mere presence of an interconnector moderates the negative effect of status that Lanz et al. (2019) report. Along these lines, we formulate the hypothesis:

H₂: As the follow-back probability of high-status influencers is much lower compared with low-status influencers, we expect that the mere presence of an interconnector moderates this negative effect of status.

In the next sections, we examine the acceleration of natural triadic closure by means of model-free evidence. Then, we test the preceding hypotheses by using propensity-score matching and conducting a logistic regression, including various controls and moderators. We also conduct an online experiment to uncover possible explanations for the formulated hypotheses and the observed results in the secondary data. Afterward, to investigate to what extent this phenomenon can support seeding efforts, we compare—by means of data-based simulations—the different seeding policies' effectiveness over time, while also considering the status quo targeting (the chosen seeding policy of the unknown creators in our data). Finally, to obtain convergent validity of our empirical findings, we conduct a preregistered field experiment.

Comparing the Effectiveness of Seeding Policies: Empirical Evidence

Data

A worldwide leading audio platform provided us with two longitudinal data sets that include approximately 35,000 users each from two different sign-up time periods: one (i.e., Data Set 1) with 24,020 content creators and 11,936 fans, and another

(i.e., Data Set 2) with 4,978 content creators and 30,022 fans. These 35,956 and 35,000 users—who all signed up in a window of one quarter (Data Set 1) and one week (Data Set 2)—were tracked from their sign-up over a period of 280 and 123 weeks, respectively (see Table W.1 in the Web Appendix for some descriptive statistics for both data sets).²

On the considered platform, there are different types of content (i.e., audio) creators such as music artists, comedians, and other podcasters, and they see the platform as an incubator because it is a unique breeding ground, especially for up-and-coming creators, to build and expand their follower base to generate more plays on their newly uploaded songs and podcasts. In short: the more followers, the greater the reach, because all followers get notified once there is a new upload.

Hence, at the beginning of the journey stands the accumulation of followers—which surveys revealed to be the primary goal of up-and-coming audio creators—and for this reason such creators choose the considered platform, as almost all others are simple streaming platforms. For example, on Spotify, Apple Music, Tidal, and Amazon Music Unlimited, unknown audio creators cannot direct outbound activities to other users of the platform, which would allow them to build a follower base and thus achieve their primary goal.

A direct analogy to the considered platform would be X, a directed network with the same features: there is a follow functionality, and content can be commented on, liked, and reposted. Naturally, the one difference is the content itself. X revolves around text, whereas the considered platform is all about audio.

Our secondary data are rich insofar as they feature time stamps for every outbound activity while capturing the formation of not only the first-degree follower base but also the second-degree follower base (which makes it a dynamic rather than static consideration). This richness allows us to study the acceleration of natural triadic closure. In the hope of getting a (direct and indirect) return, an unknown audio creator can direct an outbound activity—that is, a follow (in 73.2% of the cases), comment (6.8%), like (14.2%), or private message (5.8%)—to a user to catch their attention. Interestingly, we find that unknown audio creators direct only 11% of their outbound activities to nearby influencers, which hints at an untapped potential to improve seeding efforts.

Analysis

We consider a 2 (nearby vs. remote) × 2 (low vs. high status) setting with four seeding targets: (1) nearby low-status influencers, (2) remote low-status influencers, (3) nearby high-status influencers, and (4) remote high-status influencers.

Theoretically, nearby low-status influencers should feature the highest a priori follow-back probability: not only do we expect the presence of an interconnector to increase the follow-back probability (H_{1a}), but we also expect the number of

² Note that content (i.e., audio) creators are users who have uploaded at least one audio file, that is, one song or podcast.

interconnectors to further increase it (H_{1b}). While there is a negative effect of status (as in Lanz et al. [2019])—that is, the higher the seeding-target status, the lower the follow-back probability—we expect that the presence of an interconnector moderates this negative effect of status (H_2).

We define unknown creators according to their network status or follower base (i.e., indegree; Wasserman and Faust 1994) and apply the order-of-magnitude rationale, because the logarithmic scale is consistent with the dispersion in the number of followers on mature user-generated content platforms. In this spirit, unknown creators are those users who have less than or equal to 100 followers. We apply the same rationale to differentiate between seeding targets: Type 1 seeding targets are users with 0 to 100 followers, Type 2 seeding targets are users with 101 to 1,000 followers, Type 3 seeding targets are users with 1,001 to 10,000 followers, and Type 4 seeding targets are users with more than 10,000 followers. According to this differentiation, low-status influencers are Type 1 (≤ 100 followers) and high-status influencers are Type 4 ($> 10,000$ followers), because these two seeding policies simply look at the respective tail of the follower-base distribution. Since we define unknown creators as those with less than or equal to 100 followers, they are also Type 1, just like the low-status influencers.

Model-free evidence on the acceleration of natural triadic closure.

Concerning the direct return, Table 2 shows, for unknown (i.e., Type 1) creators, the a priori follow-back probability when targeting a low-status influencer or a high-status influencer with an outbound activity (i.e., a follow, comment, like, or private message), where the period in which we consider reciprocity in the form of follow-backs—after observing an outbound activity—was set to one week (the average login frequency). To examine natural triadic closure and the possibility to accelerate it, we calculate these a priori follow-back probabilities dependent on the absence/presence of one or several interconnectors—observed in the data.

If a seeding target is not interconnected with the unknown creator, and interconnectors are thus absent (first column), then we remain in the current seeding paradigm with regard to targeting remote low-status and high-status influencers. To move beyond it and take into account the closing of triangles, Table 2 further shows the a priori follow-back probabilities if the number of interconnectors is more than one (second column), more than two (third column), or more than three (fourth column). For example, the presence of more than one interconnector in the second column means that the unknown creator and the seeding target are interconnected via at least one user, or interconnector.

As expected, we find that *remote low-status influencers* feature the highest a priori follow-back probability (6.68%; first column, first row), and this probability is strictly decreasing with increasing status (first column, first to fourth rows; $p < .001$ for the three differences), where *remote high-status influencers* feature the lowest a priori follow-back probability (.02%; first column, fourth row). When taking into account the presence of an interconnector, we find that the a priori follow-back

Table 2. Follow-Back Probabilities and Repost Probabilities for Type 1 Creators, Dependent on the Number of Interconnectors.

	Number of Interconnectors			
	None	≥ 1	≥ 2	≥ 3
A: Follow-Back Probabilities				
Type 1 seeding target	6.680%	14.368%	17.591%	19.778%
Type 2 seeding target	2.222%	9.099%	11.399%	12.876%
Type 3 seeding target	.221%	2.004%	3.046%	3.886%
Type 4 seeding target	.018%	.185%	.044%	.000%
B: Repost Probabilities				
Type 1 seeding target	.050%	.096%	.141%	.230%
Type 2 seeding target	.029%	.022%	.018%	.016%
Type 3 seeding target	.002%	.020%	.016%	.030%
Type 4 seeding target	.001%	.000%	.000%	.000%

Notes: N (from left to right in Panel A) = Type 1: 591,795, 68,400, 31,738, and 18,076; Type 2: 511,822, 96,530, 49,471, and 30,266; Type 3: 355,888, 32,634, 13,461, and 7,179; Type 4: 148,361, 5,945, 2,257, and 931 follows, comments, likes, and messages. N (from left to right in Panel B) = Type 1: 96,116, 13,532, 5,673, and 3,050; Type 2: 101,575, 23,310, 10,861, and 6,373; Type 3: 136,669, 15,007, 6,189, and 3,324; Type 4: 111,095, 2,759, 769, and 316 follows, comments, likes, and messages. Reaction period = 1 week.

probability of *nearby low-status influencers* more than doubles, from 6.68% to 14.37% (second column, first row; $p < .001$), which provides support for H_{1a} . In contrast, targeting high-status influencers is associated with a much lower probability (.02%; first column, fourth row), even when they are second-degree followers and thus *nearby high-status influencers* (.19%; second column, fourth row). Therefore, we conclude that in our 2×2 setting, low status dominates high status, and nearby dominates remote, such that targeting nearby low-status influencers is the dominant seeding policy.

In fact, the more interconnectors present between the unknown creator and the low-status influencer, the higher the expected direct return in the form of follow-backs. Note that this monotonicity is concave, as the probabilities increase from 6.68% to 14.37%, 17.59%, and 19.78%, whereas the marginal increase in probability decays from 7.69 to 3.22 and 2.19 percentage points, respectively. These differences (i.e., from 14.37% to 17.59% as well as from 17.59% to 19.78%) are significant ($p < .001$), supporting H_{1b} . Hence, the presence of at least one interconnector when targeting low-status influencers makes all the difference (while the number adds to it). This finding speaks for the possibility of accelerating natural triadic closure, and supports the conjecture that targeting nearby low-status influencers is the dominant seeding policy among the four we are considering.

Can we draw the same conclusion regarding the *indirect return* (i.e., the a priori repost probability) when targeting a low- or high-status influencer with an outbound activity? Indeed, in Table 2 we find that low-status influencers feature the highest a priori repost probability (.05%; first column, fifth row), which is by order of magnitude higher ($p < .001$) than that of high-status influencers (.001%; first column, eighth row). In line with the findings from Table 2, we also find that the presence of an interconnector makes all the

difference, increasing the expected indirect return in the form of reposts from .05% to .1% ($p < .1$).

Note that reposts from an influencer generate further follow-backs, but indirectly, as they originate from the second-degree follower base of this influencer, where the corresponding probabilities of such an indirect return are generally much lower compared with the direct return (the follow-back from the influencer directly; compare Panels A and B in Table 2). In the data-based simulations, we take into account both of these returns, and others (as well as in higher resolution).

Acceleration of natural triadic closure above and beyond homophily as measured by the Hellinger distance. As mentioned, there are two fundamental processes of natural network formation, and since we focus on accelerating natural triadic closure, the question arises whether the *mere presence* of an interconnector makes all the difference, *above and beyond homophily*, which is the other fundamental process of natural network formation.

Shared interests are the most important characterization of homophily; hence we use the Hellinger (1909) distance (which takes the normalized form of a Euclidean distance) between the creator and the seeding target to capture the overlap in terms of audio interest and account for 11 genres, where a distance of 0 means complete overlap (i.e., exact same distribution of audio interests) and 1 means no overlap at all (i.e., audio interests are distributed across completely different genres; see Web Appendix A for a detailed description). For a much more granular measure, in a second step we use the Hellinger distance to account for interest overlap on the creator level instead of just on the genre level. The rationale is that, for example, pop music may potentially consist of very different artists.

These two types of overlap (i.e., on the genre and creator levels) are, in fact, a direct measure of homophily, or similarity, which is rarely available. To this end, we apply propensity-score matching for fair comparison between the outbound activities directed from unknown (Type 1) creators to nearby low-status influencers (where an interconnector is present; treatment), and the outbound activities directed to remote low-status influencers (where an interconnector is not present; control).

We calculate scores denoting the probability of having an interconnector present, given an outbound activity, which means we are further investigating the difference between the first and second columns in the first row of Panel A of Table 2. To calculate these scores, we use extreme gradient (tree) boosting (Chen and Guestrin 2016), weighted for class imbalance, to predict for each outbound activity the presence of one interconnector. Note that we specify the evaluation metric as log loss and add a regularization term, $\gamma = 1$, which prevents overfitting. For prediction, we use a variety of variables characterizing the unknown creator and seeding target as well as the interrelation between them: (1) the Hellinger distance between the creator and the seeding target, (2) the squared Hellinger distance, (3) the seeding target's indegree (follower base), (4) the creator's outdegree, (5) the seeding target's outdegree, (6) the creator's daily average of outbound activities, and (7) the seeding target's daily average of outbound activities.

Given the scores, we apply one-to-one matching without replacement, by matching the controls to the few treatments, which creates two buckets. The outbound activities directed to nearby low-status influencers (treatment) are in one bucket, and the (propensity-score) matched outbound activities directed to remote low-status influencers (control) are in the other bucket.

In line with the observations from Table 2, we find that there is still a major difference in a priori follow-back probabilities: a two-sample two-tailed test for equality of proportions with continuity correction gives a p -value less than .001 for both Hellinger distance measures (genre and creator level). To gain more robustness of this statistical finding, we calculate the bootstrapped mean follow-back difference over 1,000 iterations, which does not include zero, namely for the distance measure on the genre level as well as on the creator level (see Figure W.1 in the Web Appendix for the two distributions). This provides further support for H_{1a} .

Concerning the presence of more than one interconnector, note that we also apply propensity-score matching for outbound activities directed to nearby low-status influencers with more than one interconnector present (treatment) and to nearby low-status influencers with only one interconnector present (control), and again find a major difference in a priori follow-back probabilities ($p < .001$, for both the genre level and the creator level). This provides further support for H_{1b} .

Controls and moderators. For more robustness of the findings from the model-free evidence and propensity-score matching, we conduct a binary logistic regression, again focusing on unknown (Type 1) creators. To uncover how the nonlinearities in the estimation impact the magnitude of the interaction between the presence of an interconnector and the continuous variables used in the logistic regression (e.g., the overlap of music interest), we calculate the marginal association of the presence of an interconnector with these variables, based on the predicted probability of getting a follow-back (as the interaction coefficient's sign could even flip; Ai and Norton 2003).

In the logistic regression, we use the same variables as before: (1) the presence of an interconnector (x_1 ; baseline is the absence of interconnectors), (2) the Hellinger distance between the creator and the seeding target to capture the overlap in music interest (x_2 ; see Web Appendix A for a detailed description), (3) the squared Hellinger distance (x_3), (4) the seeding target's indegree (x_4), (5) the creator's outdegree (x_5), (6) the seeding target's outdegree (x_6), (7) the creator's daily average of outbound activities (x_7), and (8) the seeding target's daily average of outbound activities (x_8). In addition, we use (9) the outbound activity type, that is, follow, comment, like, or private message (\vec{x}_9 ; baseline is a follow).³

³ The component x_{91} equals 1 if the outbound activity type is a comment (and 0 otherwise), x_{92} equals 1 if it is a like (and 0 otherwise), and x_{93} equals 1 if it is a private message (and 0 otherwise). Accordingly, the case of $x_{91} = x_{92} = x_{93} = 0$ refers to a follow, which is the baseline.

Prob is the probability of a follow-back, and ϕ is the score of the full model (see Model 6 in Table 3) with variables x_1 to \vec{x}_9 :⁴

$$\text{Prob} = \frac{e^\phi}{1 + e^\phi}, \quad (1)$$

where $\phi = \beta_0 + \beta_1 x_1 + \sum_{k=2}^8 \beta_k x_k + \vec{\beta}_9 \cdot \vec{x}_9 + \beta_{12} x_1 x_2 + \sum_{k=4}^8 \beta_{1k} x_1 x_k + (\vec{\beta}_9 \cdot \vec{x}_9) x_1$.

Then, in the spirit of Ai and Norton (2003), we analyze the interaction between the presence of an interconnector and the continuous variables used in the logistic regression, that is, $k = \{2, 4, 5, 6, 7, 8\}$:

$$\text{Interaction}_{1k} = \left. \frac{\partial \text{Prob}}{\partial x_k} \right|_{x_1=1} - \left. \frac{\partial \text{Prob}}{\partial x_k} \right|_{x_1=0}, \quad (2)$$

where $\frac{\partial \text{Prob}}{\partial x_k} = \frac{e^\phi}{(1+e^\phi)^2} (\beta_k + \beta_{1k} x_1)$.

For the logistic regression, Table 3 shows ten different models. Models 1 to 5 do not feature interactions, whereas Models 6 to 10 do. Note that in Models 2 to 5 and 7 to 10 we hold the outbound activity type constant: follow (Models 2 and 6), comment (Models 3 and 7), like (Models 4 and 8), and private message (Models 5 and 10).

First of all, considering Model 1, we find that the mere presence of an interconnector increases the follow-back probability ($p < .01$), which provides further support for H_{1a} . Furthermore, the lower the overlap of music interest between the unknown creator and the seeding target, the higher the follow-back probability ($p < .1$). However, adding a quadratic term, which we find to be negative ($p < .1$), implies a trade-off: it seems that if the overlap of music interest is too minimal or too great, then the follow-back probability remains low.

Second, concerning indegree, the higher the seeding target's indegree, the lower the follow-back probability ($p < .01$), because the higher-status target is less responsive, thus confirming the findings of Lanz et al. (2019).⁵

Third, concerning outdegree, the higher the seeding target's outdegree, the higher the follow-back probability ($p < .01$), because higher outdegree implies higher responsiveness, which is in line with Valsesia, Proserpio, and Nunes (2020). Naturally, the unknown creator's outdegree has a negative coefficient ($p < .01$), because higher outdegree is associated with lower status (the seeding target realizes that the creator tends to simply target [too] many).

Fourth, similar to the outdegree effect, the higher the creator's daily average of outbound activities, the less focused the seeding efforts—the creator targets (too) many and may even tend to spam—and therefore the lower the follow-back probability ($p < .01$).

In contrast, the higher the seeding target's activities, the lower the follow-back probability ($p < .01$), because while the seeding target's activities may drive the follow-back probability, it also implies that the target has more experience and may thus be pickier.

Fifth, concerning the outbound activity type, commenting on a seeding target's content increases the follow-back probability, compared with following the seeding target ($p < .01$), whereas the other two types of activity are not as effective. In particular, private messaging is somewhat surprisingly less effective, perhaps due to content dependence, which we do not observe.

In the next step, we employ the full model and use all the variables as moderators (apart from the squared Hellinger distance) and interact them with the presence of an interconnector. This full Model 6 shows that all coefficients are very similar to the partial Model 1, including their significance. Note that the quadratic term for the overlap of music interest has a p -value of .104—in Model 1 it was .092—and can thus also be considered significant at the 10% level. We also find several moderators: the seeding target's indegree positively moderates the effect of the presence of an interconnector ($p < .01$). When calculating the marginal association of the presence of an interconnector with the seeding target's indegree, based on the predicted probability of getting a follow-back, we find that for more than 99% of the observations it is indeed positive (see Panel 1 in Figure W.2 in the Web Appendix; see also Equation 2 for the calculation of the marginal association, based on the predicted probability for each observation). This implies that the mere presence of an interconnector moderates the negative effect of status. The decreasing follow-back probability with higher indegree of the seeding target is moderated if an interconnector is present (it is higher), which provides support for H_2 .

Interestingly, while the seeding target's outdegree seems to negatively moderate the effect of the presence of an interconnector ($p < .01$), when calculating the marginal association (which takes into account the nonlinearities in the estimation), we find that 93.06% of the observations are actually positive (see Panel 2 in Figure W.2 in the Web Appendix). This implies that the mere presence of an interconnector increases the marginal tendency of a seeding target to respond and follow back, which is in line with what we would expect. The same holds true for the moderating effect of the seeding target's daily average of outbound activities: 92.47% of the observations are actually positive (see Panel 3 in Figure W.2 in the Web Appendix), which implies that the mere presence of an interconnector increases the marginal tendency of an (active) seeding target to respond and follow back.

The positive moderating effect of both the unknown creator's outdegree and this creator's daily average of outbound activities can be attributed to a by-product of the negative association of outdegree and outbound activities with the follow-back probability. More precisely, as an increase in the outdegree and/or outbound activities is associated with a lower follow-back probability—due to spamming perceptions as described previously—the mere presence of the interconnector is expected to moderate this perception in the eyes of the seeding targets.

These two flips of coefficient signs underline the need to account for nonlinearities in the estimation of logistic models.

⁴ In our subsequent analysis, we also consider partial models, that is, without interactions, as well as separate models in which we hold the outbound activity type constant, and hence the variable \vec{x}_9 is not meaningful.

⁵ Note that we focus on unknown (Type 1) creators; hence we did not use the status of the creator as a variable.

Table 3. Output of Binary Logistic Regressions Without Interactions (Models 1 to 5) and with Interactions (Models 6 to 10), and with Fixed Outbound Activity Type: Follow (Models 2 and 6), Comment (Models 3 and 7), Like (Models 4 and 8), and Private Message (Models 5 and 10).

	Dependent Variable: Follow-Back									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Interconnector presence	.525*** (.042)	.539*** (.048)	.350*** (.129)	.800*** (.146)	.206 (.186)	.670*** (.187)	.408* (.216)	.319 (.498)	1.775*** (.639)	1.345* (.783)
Hellinger distance	.515* (.267)	.799** (.311)	-1.534* (.835)	.644 (.874)	-.002 (1.127)	.510* (.272)	.821*** (.317)	-1.854** (.853)	.879 (.915)	.032 (1.144)
Squared Hellinger distance	-.500* (.297)	-.768** (.345)	1.319 (.939)	-.919 (.940)	.468 (1.287)	-.484 (.298)	-.774** (.346)	1.517 (.948)	-1.093 (.960)	.500 (1.292)
Seeding target's indegree	-.874*** (.012)	-.915*** (.015)	-.832*** (.037)	-.883*** (.039)	-.800*** (.045)	-.894*** (.013)	-.940*** (.016)	-.873*** (.041)	-.861*** (.042)	-.811*** (.048)
Creator's outdegree	-.274*** (.010)	-.368*** (.012)	-.115*** (.029)	-.084** (.039)	-.026 (.031)	-.276*** (.011)	-.374*** (.013)	-.109*** (.032)	-.085** (.041)	-.026 (.032)
Seeding target's outdegree	.692*** (.014)	.726*** (.016)	.652*** (.040)	.730*** (.049)	.645*** (.045)	.706*** (.014)	.741*** (.017)	.690*** (.043)	.707*** (.051)	.661*** (.047)
Creator's activities	-.218*** (.036)	-.138*** (.036)	-.126 (.161)	-.607** (.293)	-.483** (.206)	-.232*** (.040)	-.159*** (.041)	-.187 (.181)	-.143 (.270)	-.334 (.205)
Seeding target's activities	-.166*** (.028)	-.135*** (.033)	-.346*** (.086)	-.251** (.107)	-.247** (.096)	-.107*** (.031)	-.082** (.036)	-.338*** (.099)	-.038 (.116)	-.239** (.101)
Comment	.201*** (.051)					.223*** (.056)				
Like	-.294*** (.058)					-.354*** (.064)				
Private message	-.294*** (.061)					-.278*** (.065)				
Interconnector × Hellinger distance						.005	-.057	1.072	-.319	-.533
Interconnector × seeding target's indegree						(.217)	(.247)	(.703)	(.778)	(1.159)
Interconnector × creator's outdegree						.169*** (.036)	.218*** (.043)	.263*** (.095)	-.184 (.127)	.162 (.148)
Interconnector × seeding target's outdegree						.025	.064*	-.037	.268*	.026
Interconnector × seeding target's outdegree						(.029)	(.036)	(.084)	(.148)	(.121)
Interconnector × creator's activities						-.182*** (.045)	-.199*** (.054)	-.336*** (.128)	.107 (.168)	-.308 (.192)
Interconnector × seeding target's activities						.072 (.088)	.086 (.088)	.239 (.391)	-7.670*** (2.018)	-2.456** (1.139)
Interconnector × seeding target's activities						-.190** (.099)	-.162* (.162)	.099	-.870** (.186)	.186

(continued)

Table 3. (continued)

	Dependent Variable: Follow-Back									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Interconnector × comment						(.075)	(.088)	(.204)	(.350)	(.343)
						-.163				
						(.139)				
Interconnector × like						.318**				
						(.150)				
						-.177				
						(.193)				
Constant	-.947***	-.657***	-.638***	-1.880***	-1.722***	-.945***	-.627***	-.540**	-2.106***	-1.786***
	(.073)	(.085)	(.203)	(.238)	(.257)	(.078)	(.091)	(.213)	(.259)	(.266)
Log-likelihood	-17,557.230	-12,775.870	-1,830.008	-1,482.624	-1,350.074	-17,523.020	-12,752.060	-1,824.076	-1,460.331	-1,345.026
Alkaike information criterion	35,138.470	25,569.730	3,678.016	2,983.248	2,718.148	35,088.030	25,534.120	3,678.152	2,950.661	2,720.053

Note: *p < .1.

**p < .05.

***p < .01.

Similarly, for the interaction between the presence of an interconnector and the overlap in music interest, we find that although the coefficient is insignificant, more than 99% of the observations are positive (see Panel 4 in Figure W.2 in the Web Appendix). This implies that a seeding target is more likely to follow back a more distant unknown creator if there is an interconnector present. Finally, concerning the outbound activity type, we find that the mere presence of an interconnector moderates the inferiority in effectiveness of a like versus a follow.

Considering all models other than Models 1 and 6, first, we find in Model 5 that the presence of an interconnector is insignificant for private messages, perhaps due to content dependence, which we do not observe. Second, concerning Hellinger distance, the relation is either increasing and concave for follows (Models 2 and 7) or decreasing for likes (Models 3 and 8), or it is insignificant (all other models). Third, for all other variables, we find consistent signs, that is, for the seeding target's indegree (negative), the creator's outdegree (negative), the seeding target's outdegree (positive), the creator's daily average of outbound activities (negative), and the seeding target's daily average of outbound activities (negative). Finally, we find similar results for the interactions for each outbound activity type, which concern Models 7 to 10. Therefore, the findings from Model 6 replicate.

Online Experiment

We examine possible explanations for the result observed in the secondary data by means of an experimental study (note that theory development plays a supporting role in this article). We ran an exploratory survey gathering mediators to complement the one based on theory (i.e., trust; Easley and Kleinberg 2010), before running a preregistered (<https://aspredicted.org/5xa3f.pdf>) online experiment to test these mediators in the relationship between the presence/absence of an interconnector and the follow-back intention, with 500 respondents in each of the two conditions.

In terms of mediators, we consider (1) perceived trust and, based on the exploratory survey, (2) perceived similarity as well as (3) (favor-)reciprocity. We always let respondents evaluate statements on a five-point Likert scale (1 = "strongly disagree," and 5 = "strongly agree"), and upon randomizing the statements in a pilot, we did not observe an order effect.

For perceived trust, we use a measurement taken and adapted from Thomson (2006), namely "This user can be trusted," "This user can be counted on," and "This user is dependable." In an open-ended question in the exploratory survey, in which we explicitly asked respondents why they chose to follow back, respondents consistently provided similar explanations such as "I would probably be automatically trusting as I am quite selective with how I follow, hence I would see they are already connected within my network" or "I trust the people I am personally friends with, so if someone I know follows another individual I am less inclined to judge the person as threatening."

For perceived similarity (sometimes also referred to as perceived homophily; Meyners et al. 2017), we use measurements partially taken and adapted from Reis et al. (2011) and

Swaminathan et al. (2007), namely "This user is similar to me" and "This user and I have a lot in common," because they align with explanations consistently provided in the exploratory survey: "The fact that I also follow someone else who likes their content, who follows them, that's just icing on the cake. It feels more like a connection, like a group of like-minded people found each other finally," or "It would mean we probably would share some common interests."

For (favor-)reciprocity, we use a measurement partially taken and adapted from Perugini et al. (2003), which was also used in large panels (e.g., Dohmen et al. 2008, 2012), namely "This user followed me, hence I return the favor and follow back." This notion of (favor-)reciprocity goes in line with the explanation consistently provided in the exploratory survey: "I like their content and they followed me first, so I will return the favor" or "I like this person and I feel like they took initiative to signal to me that they are interested in me. Thus, I would return the favor to them." Probably (favor-)reciprocity happens due to a perceived sense of obligation, as this respondent's explanation suggests: "I believe in making connections and if this person chose to follow me, and they had some kind of relationship with someone who follows me, I would feel that it would be obvious I should follow them back."

We ran a preregistered online experiment to test these mediators. First, we find that with an interconnector present, the follow-back intention is significantly higher (Mean = 3.71), compared with when an interconnector is absent (Mean = 3.12), as a two-tailed Welch's unequal variances t-test ($p < .001$) and the bootstrapped mean scores difference (not including zero) show. Second, we find that (1) perceived trust, (2) perceived similarity, and (3) (favor-)reciprocity partially mediate the relationship between the presence/absence of an interconnector and the follow-back intention. Without the mediation, the presence of an interconnector yields a coefficient of .58 (it adds .58 to the follow-back intention score), and with the mediation it is reduced to .19, implying that the mediators explain 68% of the effect. We also find that the effects of all mediators are significant (see Figure W.3 in the Web Appendix for an overview). Note that we also run a bootstrapping analysis of the follow-back intention on the treatment via the three mediators, and we find that for both similarity and (favor-)reciprocity it does not include zero, whereas for trust 99.8% of the samples yield values greater than zero. This suggests that trust is not as strong a mediator as similarity and (favor-)reciprocity.

Data-Based Simulations

From the previous model results, we can conclude that outbound activities can accelerate natural triadic closure, as the mere presence of an interconnector increases the follow-back probability, and this also occurs when accounting for a variety of controls and moderators. But to what extent can this phenomenon support seeding efforts?

In this subsection, we compare the effectiveness of the different seeding policies over time by means of data-based simulations (1,000 iterations, or "realizations of the world").

According to the 2 (nearby/remote) \times 2 (low/high status) setting, we consider targeting (1) nearby low-status influencers, (2) remote low-status influencers, (3) nearby high-status influencers, and (4) remote high-status influencers. In addition, we consider (5) the status quo targeting—the actual chosen seeding policy of the unknown creators in our data—to put these seeding policies in perspective. The calculated and analyzed a priori probabilities for unknown creators provide the foundation, because in the data-based simulations we use the example of an audio creator who has just signed up to the platform and thus has zero followers at the beginning.

1. In the first seeding policy, the simulated creator adopts our suggested seeding policy: Upon acquiring the first follower after targeting *remote* low-status influencers (i.e., with less than or equal to 100 followers, namely Type 1 seeding targets), the creator immediately focuses on the (now existing) second-degree follower base to benefit from the substantially higher responsiveness, and targets *nearby* low-status influencers. Recall that the a priori follow-back probability doubles when leveraging first-degree followers as interconnectors to close triangles. Hence, whenever there is an available seeding target that is of low status and part of the second-degree follower base, then the simulated creator selects this target. Otherwise, the creator falls back to targeting remote low-status influencers, before targeting again the nearby low-status influencers once a further follower has been acquired.
2. In the second seeding policy, the creator directs outbound activities exclusively to remote low-status influencers.
3. In the third seeding policy, the creator directs outbound activities—if available—to *nearby* high-status influencers (i.e., with more than 10,000 followers, namely Type 4 seeding targets in the second-degree follower base). Otherwise, the simulated creator falls back to targeting remote low-status influencers.
4. In the fourth seeding policy, the creator adopts the traditional influencer marketing approach and directs outbound activities exclusively to remote high-status influencers.
5. Finally, we include a real-life benchmark, the status quo targeting, which is the unknown creators' actual seeding policy as observed in our data, taking into account not only the selection of different seeding-target types by these creators but also the division into nearby and remote ones, which creates eight buckets: Type 1 seeding targets (low-status influencers) that are nearby (3.78%) versus remote (32.67%), Type 2 seeding targets that are nearby (5.33%) versus remote (28.26%), Type 3 seeding targets that are nearby (1.80%) versus remote (19.65%), and Type 4 seeding targets (high-status influencers) that are nearby (.33%) versus remote (8.19%). This selection distribution reveals that unknown creators in our data are not

complying with any of the considered seeding policies, and the data-based simulations enable us to understand how their approach performs in comparison.

Method. We compare the seeding policies' effectiveness over the course of 24 months by calculating the median growth of the follower base. In each month, the simulated creator directs 40 outbound activities to the respective seeding targets (i.e., 10 per week). Note that with these outbound activities, a user cannot be targeted twice. Consequently, over the course of 24 months a creator selects 960 different seeding targets according to the respective seeding policy.

In the data-based simulations, given an outbound activity by a creator, we take into account the return on a seeding target—whether it is a direct or indirect one—in high resolution concerning creator status, seeding-target status, and the number of interconnectors present (if any). We also consider cascades, as there are long-term, additional indirect returns on a newly acquired (first-degree) follower (i.e., the seeding target who followed back), namely in the form of follow-backs from the second- and third-degree follower base due to future reposts by this newly acquired (first-degree) follower and that follower's followers. Therefore, the data-based simulations emulate rich network dynamics and reveal the long-term impacts of each seeding policy.

Each creator initially has zero followers and is therefore a Type 1 creator. Unlike in Table 2, in the data-based simulations we update the status-dependent probability of a direct and indirect return—which are observed in the data—in multiples of 25 followers (i.e., the resolution is higher). More precisely, the creator receives improved probabilities after achieving a follower base of 25, then after achieving a follower base of 50, and so forth.

Naturally, it is possible that the creator gains only an indirect return, without a direct one (i.e., the seeding target did not follow back). In this case, there is no additional indirect return, because of the lack of connection. The seeding target simply reposts a song or podcast from the creator but does not follow back and is therefore not part of the creator's first-degree follower base (in which the seeding target could be exposed to the creator's future uploads).

Along these lines, since a user cannot be targeted twice, there are four scenarios: (1) a seeding target does not follow back (i.e., no return), (2) a seeding target follows back and potentially reposts in the future (i.e., a direct return plus potential additional indirect returns), (3) a seeding target follows back and reposts as well as potentially reposts in the future (i.e., a direct return plus an indirect return plus potential additional indirect returns), or (4) a seeding target only reposts (i.e., only an indirect return). In Web Appendix B, we provide a detailed description of the different returns we account for in the data-based simulations.

Independent of the seeding policy, each simulated creator also acquires followers naturally, that is, without targeting users with outbound activities. Hence, besides the observed probabilities of the different returns from a seeding target, there is a probability of baseline follows—whose status-dependent distributions are also observed in our data—which

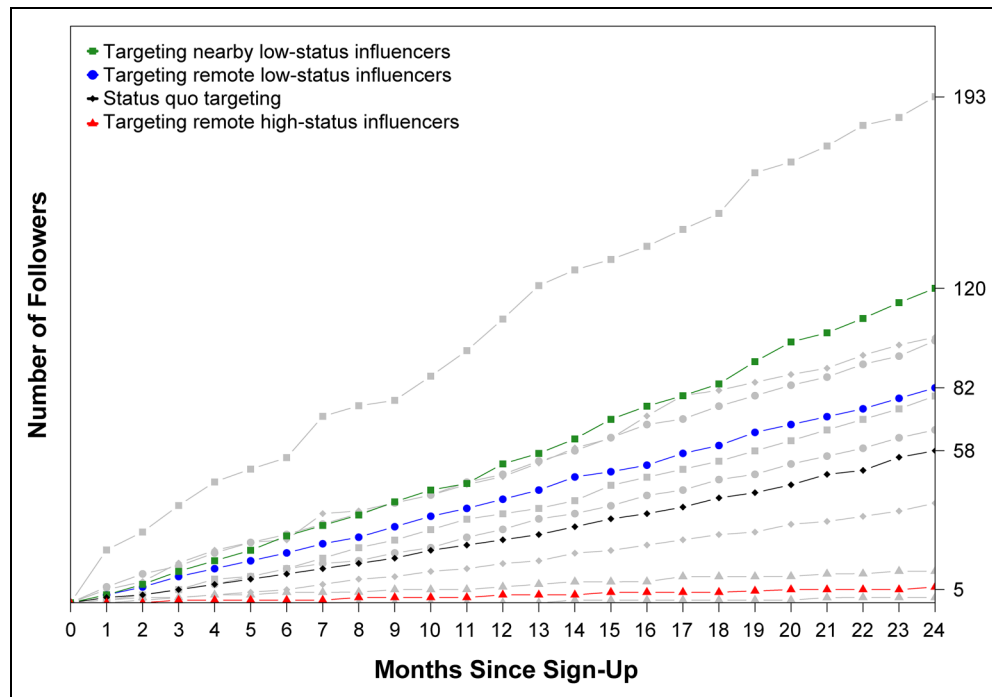


Figure 2. The Median Growth of the Follower Base with a Budget of 40 Monthly Outbound Activities, Including the Boundaries Given by 90% Confidence Intervals.

build and expand the creator's follower base. Recall that typically 48% of this baseline originates in the second-degree follower base, and 52% comes from all over the platform.

Finally, in our data there are 35,956 users (Data Set 1); hence we have 35,956 egocentric networks that reach two degrees out, consisting of a total of 32,404,028 users. Thereby we have a proxy of the network structure of the whole platform, which we use as the social network in our data-based simulations. Note that the density of the underlying network is low: the clustering coefficient amounts to .0001. Therefore, the network coupling effects are relatively marginal. Also note that the availability of seeding targets, which is crucial in our context, is more than sufficient, with 31,740,844 Type 1 seeding targets (low-status influencers) and 5,619 Type 4 seeding targets (high-status influencers). With this heavy-tailed follower-base distribution, the creator faces no availability constraint in any of the considered seeding policies.

Results. Figure 2 shows the median growth of the follower base with a budget of 40 monthly outbound activities, including the 90% confidence intervals. In this figure, the line with triangles represents the growth when targeting remote (and nearby) high-status influencers, the line with diamonds represents the growth of the status quo targeting, the line with circles represents the growth when targeting remote low-status influencers, and the line with squares represents the growth when targeting nearby low-status influencers.

First, we find that exclusively targeting remote high-status influencers over the course of 24 months, which is the traditional influencer marketing approach, results in a median

follower base of five followers with very narrow 90% confidence intervals. This means that due to the extremely low responsiveness of high-status influencers—they rarely follow back—the five followers practically consist only of baseline followers, namely the followers one acquires naturally and without targeting. The baseline, which is accounted for in all seeding policies (where typically half originates in the second-degree follower base and can thus be attributed to natural triadic closure), is therefore very low. This suggests that the acceleration of natural triadic closure—given that targeting nearby low-status influencers results in a median follower base of 120 (discussed subsequently)—goes far beyond what is typically achieved by natural triadic closure.

Second, the results from targeting remote high-status influencers—or trying to do so—are indistinguishable from targeting nearby high-status influencers, because unknown creators are bound to the fallback option as a consequence of the availability problem: Due to how networks are built—that is, high-status influencers typically feature a very low outdegree, and if they follow someone, they follow other high-status influencers—it is very rare that second-degree followers are of high status, especially given that we consider unknown creators. Moreover, the a priori follow-back probability of nearby high-status influencers increases from .02% to .19% (see Table 2), which is still extremely marginal. Hence, even if a nearby high-status influencer is available, effectively the return from them is not feasible.

Third, we find that exclusively targeting remote low-status influencers over the course of 24 months results in a median follower base of 82, outperforming the traditional influencer marketing approach by an order of magnitude, as expected.

Fourth, leveraging first-degree followers as interconnectors to close triangles results in a median follower base of 120. As a result, our suggested policy of targeting nearby low-status influencers is highly effective, as it outperforms the traditional influencer marketing approach by 2,300% and outperforms the approach of targeting remote low-status influencers by 46%.

Fifth, when comparing our suggested policy and the second best, especially the distributions, we find that the 95th percentile of the second best (100 followers) falls short of the median of our suggested policy (120 followers). Also note that while the second best features a symmetric distribution—90% of all observations lie between 66 and 100 followers—the distribution of our suggested policy is highly skewed: the 5th percentile (79 followers) is almost aligned with the median of the second best (82 followers); however, the 95th percentile amounts to 193 followers. More specifically, 90% of all observations lie between 79 and 193 followers, with a median of 120. Hence, 45% of all observations lie between 120 and 193 followers. Comparing means instead of medians gives even stronger results: 125 followers (our suggested policy) versus 82 followers (second best). In that case, our suggested policy outperforms targeting remote low-status influencers by 52%.

Sixth, assuming that the simulated creator is already somewhat known—say the creator has crossed an order of magnitude and initially starts with 100 followers—then leveraging first-degree followers as interconnectors to close triangles results in a median follower base of 247, instead of 120 when initially having zero followers. The factor of $\frac{247-100}{120-0} = 1.23 > 1$ suggests that with an increasing number of followers, the slope of the growth gets steeper, which implies that the convexity starts to take effect, whereby the input (outbound activities) results in growing output (follow-backs) with increasing network status. The 95th percentile amounts to 309 instead of 193—and thus the factor is also above 1, that is, $\frac{309-100}{193-0} = 1.08$. Note, however, that with increasing network status, the objective function of creators may change (e.g., the focus may shift to keeping the links to the follower base active).

Finally, for the real-life benchmark, which is the unknown creators' actual seeding policy as observed in our data, we find that noncompliance with the seeding policy of targeting either remote or (even better) nearby low-status influencers allows these creators to outperform only the traditional influencer marketing approach: the median follower base of the status quo targeting amounts to 58, and 90% of all observations lie between 38 and 101 followers.

Therefore, in line with the findings from the model-free evidence, we conclude that in our 2 (nearby/remote) \times 2 (low/high status) setting, the two dominant seeding policies revolve around targeting low-status influencers, with nearby being the most dominant (by 46%).

Comparing the Effectiveness of Seeding Policies: Experimental Evidence

Secondary data by definition involve self-selection in directing outbound activities to seeding targets, and thus a potential missing-variable problem may arise in the analysis. In this section, to

obtain convergent validity of our findings in a (more) controlled environment, we replicate the comparison of the two dominant seeding policies—which revolve around targeting low-status influencers—with a preregistered (<https://aspredicted.org/f364y.pdf>) field experiment on the same platform, investigating the effectiveness of targeting *nearby* low-status influencers (the most dominant), who are benchmarked against *remote* low-status influencers (Experiment 1). Given the widespread support and application of the traditional influencer marketing approach, we also benchmark them against remote high-status influencers as an additional check (Experiment 2).⁶

Data

In line with the previous study, we focus on audio creators who have just signed up on the platform, and we direct ten outbound activities to the respective seeding targets on a weekly basis, over an eight-week period. Here, we consider audio curators, a type of audio creator (similar to DJs). Following this rationale, we created 30 new profiles for each seeding policy and evenly assigned them to the two main music genres on the platform: 15 new profiles in the “Hip-hop & Rap” domain and 15 new profiles in the “Electronic” domain (for each seeding policy). Both Experiment 1 (which benchmarks targeting nearby low-status influencers against targeting remote low-status influencers) and Experiment 2 (which benchmarks targeting nearby low-status influencers against targeting remote high-status influencers) cover a total of 60 profiles each, where 30 are in the “Hip-hop & Rap” domain and 30 are in the “Electronic” domain (evenly divided among the two seeding policies). We choose the sample size of 30 because under these conditions the pattern of the binomial distribution is similar to that of a normal distribution (as approximated using the central limit theorem). In total, we created $2 \times 60 = 120$ new profiles.

Upon creation, each profile received a profile name, based on the 75 most popular male and female names in the United States in 2019,⁷ and five random numbers, for example, Liam79661. We randomly assigned these profiles to a genre and seeding policy, where in a given genre (i.e., “Hip-hop & Rap” or “Electronic”) each profile received the same profile picture (see Figures W.4 and W.5 in the Web Appendix for the two pictures) and, since we consider audio curators, a random playlist entitled “My favourites” consisting of 20 songs from audio creators who were trending at the time of creation of the respective profile (see Figure W.6 in the Web Appendix for a screenshot of an exemplary playlist).

In Experiment 1, 30 creators (15 in the “Hip-hop & Rap” domain and 15 in the “Electronic” domain) direct outbound

⁶ We split the field experiment into two due to logistical reasons; each seeding policy required intense work of 6 students over two consecutive eight-week periods: 12 students managed 60 profiles for Experiment 1, and these 12 students then managed another 60 profiles for Experiment 2.

⁷ We obtained the list from <https://www.ssa.gov/OACT/babynames/index.html> (accessed February 19, 2021).

activities in the form of follows only to remote low-status influencers, that is, actual users with less than or equal to 100 followers (Type 1 seeding targets). To ensure activity of the seeding target and the fit between the 30 creators as well as the seeding targets, we randomly selected as remote low-status influencers users who reposted a trending song (also randomly selected) in the respective genre (see Figure W.7 in the Web Appendix for screenshots of the clickstream to a remote low-status influencer).

The other 30 creators (15 in the “Hip-hop & Rap” domain and 15 in the “Electronic” domain) adopt our suggested seeding policy: whenever available, they direct outbound activities in the form of follows to the second-degree follower base (see Figure W.8 in the Web Appendix for screenshots of the clickstream to a nearby low-status influencer). Otherwise, the 30 creators fall back to targeting remote low-status influencers before targeting again the nearby low-status influencers once further followers have been acquired.

In Experiment 2, also covering a total of 60 profiles, we benchmark targeting nearby low-status influencers against targeting remote high-status influencers. Therefore, in the benchmark policy, 30 creators (15 in the “Hip-hop & Rap” domain and 15 in the “Electronic” domain) adopt the traditional influencer marketing approach and direct outbound activities in the form of follows only to remote high-status influencers, that is, actual users with more than 10,000 followers (Type 4 seeding targets). We randomly selected remote high-status influencers among the trending audio creators in the respective genre, as well as in the outdegree⁸ of such high-status influencers (see Figure W.9 in the Web Appendix for screenshots of the clickstream to a remote high-status influencer). The other 30 creators (15 in the “Hip-hop & Rap” domain and 15 in the “Electronic” domain) adopt our suggested seeding policy as described in Experiment 1.

Method

For both Experiments 1 and 2, and for each of the two seeding policies in them, we calculate—in line with the analysis of the data-based simulations—the median growth of the follower base over an eight-week period. In each week, starting from April 5, 2021 (Experiment 1), and June 7, 2021 (Experiment 2), the creators direct ten outbound activities in the form of follows to the respective seeding targets. Thus, over the course of the eight weeks, a creator selects 80 different seeding targets according to the respective seeding policy. To ensure independence among creators, we also made certain that a seeding target was not followed by more than one creator in each experiment.

⁸ Since high-status influencers are relatively rare, and since each creator needs to find $8 \times 10 = 80$ different high-status influencers over the course of eight weeks, we extend the search for (Type 4) seeding targets to include the outdegree of high-status influencers, because high-status influencers typically follow each other and have similar characteristics due to homophily.

Results

For all seeding policies we observe no reposts, and thus no indirect returns. Hence the median growth of the follower bases consists only of direct returns because of outbound activities in the form of follows. This is in line with the observed repost probabilities, which are much lower (compare Panels A and B in Table 2).

In Experiment 1, we find that targeting remote low-status influencers results in a median follower base after eight weeks of only 6 followers (Mean = 5.67), compared with 10 followers (Mean = 13.41) when leveraging the first-degree followers as interconnectors to close triangles. In Experiment 2, the traditional influencer marketing approach results in a median follower base of only 5 followers (Mean = 5.13), compared with 7 followers (Mean = 9.93) when following our suggested seeding policy.⁹

Based on a two-tailed Welch’s unequal variances t-test we find that these differences after the eight-week period are already significant ($p < .001$ in Experiment 1; $p < .01$ in Experiment 2). To account for the distributions being not fully normal, we also run a Wilcoxon rank sum test with continuity correction ($p < .001$ in Experiment 1; $p < .05$ in Experiment 2).

To gain robustness of these statistical findings, we calculate the bootstrapped mean follower base difference over 1,000 iterations. For Experiments 1 and 2, the bootstrapped mean follower base difference does not include zero (see Figure W.10 in the Web Appendix for the two distributions).

Interestingly, when creators target remote high-status influencers (Experiment 2), the baseline followers (i.e., the followers one acquires naturally and without targeting) are slightly higher due to the short gain in visibility when a creator is initially listed as a follower of a high-status influencer: other creators are searching for seeding targets among the recent followers of high-status influencers; hence the 30 creators that adopt this seeding policy achieve a higher median follower base after the eight-week period (but this policy is still largely inferior, as the testing for differences shows).

For a more stringent analysis, we replicate the same analysis while omitting the baseline followers. In this case, the median follower bases after eight weeks amount to 4 (Mean = 4.00) versus 7 (Mean = 9.97) followers in Experiment 1, and 4 (Mean = 4.27) versus 6 (Mean = 8.03) followers in Experiment 2, where again our suggested seeding policy dominates. Also, we find the differences to be significant, based on a two-tailed Welch’s unequal variances t-test ($p < .001$ in Experiment 1; $p < .01$ in Experiment 2). To account for the distributions being not fully normal, we also run a Wilcoxon rank sum test with continuity correction ($p < .001$ in Experiment 1; $p = .0592$ in Experiment 2).

⁹ Note that targeting nearby low-status influencers in Experiments 1 and 2 does not result in a significant difference in follower base as expected. Considering the median follower bases after eight weeks of our suggested seeding policy—10 followers (Mean = 13.41) in Experiment 1 and 7 followers (Mean = 9.93) in Experiment 2—a two-tailed Welch’s unequal variances t-test gives a p -value of .16. We also run a Wilcoxon rank sum test with continuity correction, which gives a p -value of .11. Hence, we do not find a significant difference, as we actually follow the same policy.

These findings, which may imply causality, lead to the same conclusion as the empirical evidence concerning the core role of interconnectors and the extent to which they impact seeding efforts. Due to the random assignment of profiles to seeding targets according to the respective policy, these findings provide further support for H_{1a} , namely that above and beyond the a priori matching between an unknown creator and a low-status influencer (i.e., due to homophily), the mere presence of an interconnector can make all the difference.

Implications

Although this article's contribution is primarily substantive, we begin this section by pointing out the two main theoretical implications. Concerning managerial implications, the acceleration of natural triadic closure can be applied to many different instances, and not only individuals and firms but also user-generated content platforms can implement measures accordingly, which we outline following the theoretical implications.

Theoretical

In terms of theoretical implications, first, we provide a new angle on triadic closure, which is a well-established social dynamic, found to be at play on the interorganizational and the interpersonal level (e.g., Granovetter 1973; Rajkumar et al. 2022; Schmitt, Skiera, and Van den Bulte 2011; Van den Bulte et al. 2018). We focus on the latter and present in this article a pure online setting revolving around unknown creators (individuals or firms) and seeding—a rarely studied instance due to obvious data-acquisition challenges—and investigate it with secondary data as well as a preregistered field experiment. The considered setting provides a new angle on triadic closure, because unlike in other settings such as customer referrals (e.g., Van den Bulte et al. 2018), here the interconnector—that is, the first-degree follower (or referring customer)—has a passive role. For this reason, drivers of the triadic-closure phenomenon, such as the opportunity to meet and the incentive for introductions (Easley and Kleinberg 2010), do not play a role. Note, however, that in our experiment we find that trust is indeed a driver, as previous work has pointed out, and is evident in this pure online setting. Here, we uncover two additional underlying drivers: similarity and (favor-)reciprocity.

Second, in terms of theoretical implications, we provide further empirical support for the effectiveness of targeting *low-status influencers*, a policy that seeding literature in marketing has not yet fully recognized (there are currently only a few works on it; e.g., Ameri, Honka, and Xie 2023; Beichert et al. 2023; Lanz et al. 2019). In addition, we add a new orthogonal dimension to targeting such low-status influencers, namely by shifting the focus to the follower base of the focal unknown creator to find the nearby low-status influencers, which has been previously ignored and which is conceptually different from Feld's (1991) friendship paradox that has been utilized to identify and target remote high-status influencers (e.g.,

Chin, Eckles, and Ugander 2022; Kumar and Sudhir 2021; Stephen and Lehmann 2016).

Managerial

Seeding is, in fact, at the core of user-generated content platforms: the fundamental idea is to provide the means to leverage social capital, where social capital is referred to as the value embedded in a follower base (e.g., Coleman 1988). Leveraging social capital is strongly interlinked with selecting corresponding seeding targets—within a follower base or outside—and platforms from SoundCloud to X to Instagram to LinkedIn cater to a variety of objectives: extending the fan community, acquiring new customers, or landing a new job, for example. On these platforms, outbound activities provide the foundation for leveraging social capital and achieving objectives.

Concerning the similarities and differences in the effectiveness of outbound activities, the functionalities are very similar across user-generated content platforms (i.e., following others, commenting on others' content and/or liking it, and sending others private messages), and the difference typically lies in the content type being uploaded on profiles (e.g., audio on SoundCloud, text on X, visuals on Instagram, professional content on LinkedIn). For the considered worldwide leading audio platform and the purpose of extending the fan community, we find that likes and follows are the most effective outbound activities to trigger follow-backs; namely, the mere presence of an interconnector is most effective if the outbound activity type is either a like or a follow (based on the sizes of the effects in Models 2 to 5; see Table 3). It should be noted that this could be due to content dependence, which we do not observe. Nevertheless, since likes and follows by definition exhibit a positive nature—while comments and private messages may take any valence and can also be (mis)interpreted in various ways—the safest and most straightforward way to build a follower base by leveraging interconnectors is to like audio uploaded by second-degree followers, which is, however, only possible if the seeding target is also a creator. An even more straightforward way is to simply follow second-degree followers, because everyone (not just creators) can be followed.

For other platforms, it could very well be different. For example, on work-related platforms such as LinkedIn, if a user wants to learn about new job opportunities, or land a new job, communicating intentions may be crucial; hence comments, and even more so private messages, can be expected to be more effective. Certainly, employees who have just been laid off should leverage interconnectors, independent of the outbound activity type. In this context, weak ties have been found to help job seekers land new employment (Rajkumar et al. 2022); however, what if there is no direct connection to the future dream job? This is when job seekers may take advantage of indirect connections, reach out to them, and let them know that they have mutual friends or acquaintances.

Also, salespeople who are using work-related platforms such as LinkedIn as a means for new customer acquisition should leverage interconnectors for access to companies that they want to sell their services to. In this context, Weiler et al. (2022) find for a LinkedIn-like platform that a commitment to using the active networking features—among others, the outbound activities that we consider in this article—directly serves the main objective of building a follower base and acquiring new customers. However, these networking features would be even more effective if there is an interconnector present.

New customer acquisition or a new employment can be directly supported by user-generated content platforms, particularly the work-related ones. Since recommender systems on such platforms predominantly focus on users connecting with one another, without accounting for their main objective, work-related platforms should, first, better understand the objective function of different users and, second, adjust their recommender systems accordingly and provide a selection of seeding targets, following the findings in this article. They should also give some direct advice in the form of best practices for users who are drafting posts, for example, the use of “new job” or related terms by job seekers or terms such as “solution” by users who want to acquire new customers.

The same holds true for the considered platform: given that up-and-comers see the platform as a unique breeding ground to build and expand their follower base to generate more plays of their newly uploaded podcasts and songs, why not offer a recommender system purely for creators, where everything is about discovery of seeding targets, as opposed to content discovery, which caters only to the needs of the fans? The large user base may enable fine-tuning of an algorithm based on our findings, while capturing further heterogeneity (e.g., tracking geographic location). At the core of this algorithm should be the leveraging of interconnectors and the advice to creators to like audio uploaded by second-degree followers, or simply to follow them.

Discussion

In this article we consider unpaid influencer endorsements, also known as the follow-for-follow approach, an emerging research branch of the seeding literature in marketing due to the circumstances of small and medium-sized businesses. In contrast to large corporations with a strong brand image, small and medium-sized businesses face a range of constraints, including financial limitations, and therefore cannot always offer compensation for endorsements. In fact, such relatively unknown content creators—whether they are individuals or firms—typically rely on the basic form of influencer marketing. This refers to the instance in which the individual or firm utilizes outbound activities—such as follows, comments, likes, and private messages—to get the influencer’s attention. In the context of such unpaid influencer endorsements, we provide both empirical and experimental evidence that the way networks are naturally built—via the closing of triangles—is most effective for building a follower base. This means that a follower base should not be built exogenously by cherry-picking remote

influencers, but rather endogenously and along the follower base of an individual or firm.

Our findings from unpaid endorsements may also have managerial implications for paid endorsements in that the firm, when selecting influencers, should first consider influencers who are (inter)connected with the firm itself—supposedly increasing the contract-acceptance rate and revenues, which is left for future research. Future research should also consider individuals or firms that have already established a platform presence and are thus known content creators, who may have different objective functions and even monetary means. It can be expected that the objective function of such known creators may not predominantly be just to expand their follower base; rather they might focus on keeping the links to the follower base active. Future research should also consider—instead of random selection of nearby low-status influencers—a strategic selection that includes subtypes of such influencers.

Acknowledgments

The authors thank Yvan Norotte, the student group that conducted the field experiment (i.e., group leader Marc Tolksdorf, Maja Peters, Pirus Schmidt, Lennart Reissner, Ivonne Mangels, Trixie Park, Nathalie Klein, Anamarija Rakojkova, Xinyi Liu, Anna Wildhirt, Xi Yu, Lukas Schreye, and Yuqing Du, as well as Teresa Eckardt and Annika Hauser, who conducted the pilot), the faculty and deanship of HEC Paris, and the seminar participants at Goethe University, Bocconi University, University of Innsbruck, Erasmus School of Economics, University of Zurich, ESMT Berlin, University College London, Imperial College London, University of Cologne, University of Basel, University of St.Gallen, Ben-Gurion University, University of Geneva, BI Norwegian Business School, Copenhagen Business School, IESE Barcelona, Católica Lisbon, EPFL Lausanne, the Marketing and the Creator Economy conference, the Choice Symposium, the Marketing Science Conference, the EMAC Conference, the SAMS Conference, and Social@IDC for valuable comments on previous versions of this article. The article greatly benefited from the comments of the entire review team.

Associate Editor

Michael Trusov


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Israel Science Foundation (grant number 1989/20), the Julius Paul Stiegler Memorial Foundation, the DAAD, and the IPID4all program.

ORCID iD

Andreas Lanz  <https://orcid.org/0000-0003-2643-1960>

References

- Ai, Chunrong and Edward C. Norton (2003), "Interaction Terms in Logit and Probit Models," *Economic Letters*, 80 (1), 123–29.
- Ameri, Mina, Elisabeth Honka, and Ying Xie (2023), "From Strangers to Friends: Tie Formations and Online Activities in an Evolving Social Network," *Journal of Marketing Research*, 60 (2), 329–54.
- Ansari, Asim, Florian Stahl, Mark Heitmann, and Lucas Bremer (2018), "Building a Social Network for Success," *Journal of Marketing Research*, 55 (3), 321–38.
- Beichert, Maximilian, Andreas Bayerl, Jacob Goldenberg, and Andreas Lanz (2023), "Revenue Generation Through Influencer Marketing," *Journal of Marketing* (published online November 16, 2023), <https://doi.org/10.1177/00222429231217471>.
- Chen, Tianqi and Carlos Guestrin (2016), "Xgboost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, 785–94.
- Chin, Alex, Dean Eckles, and Johan Ugander (2022), "Evaluating Stochastic Seeding Strategies in Networks," *Management Science*, 68 (3), 1714–36.
- Coleman, James S. (1988), "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94 (Suppl.), S95–S120.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2008), "Representative Trust and Reciprocity: Prevalence and Determinants," *Economic Inquiry*, 46 (1), 84–90.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2012), "The Intergenerational Transmission of Risk and Trust Attitudes," *Review of Economic Studies*, 79 (2), 645–77.
- Easley, David and Jon Kleinberg (2010), *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.
- Feld, Scott L. (1991), "Why Your Friends Have More Friends Than You Do," *American Journal of Sociology*, 96 (6), 1464–77.
- Galeotti, Andrea and Sanjeev Goyal (2009), "Influencing the Influencers: A Theory of Strategic Diffusion," *RAND Journal of Economics*, 40 (3), 509–32.
- Gelper, Sarah, Ralf van der Lans, and Gerrit van Bruggen (2021), "Competition for Attention in Online Social Networks: Implications for Seeding Strategies," *Management Science*, 67 (2), 1026–47.
- Goldenberg, Jacob, Sangman Han, Donald R. Lehmann, and Jae Weon Hong (2009), "The Role of Hubs in the Adoption Process," *Journal of Marketing*, 73 (2), 1–13.
- Goldenberg, Jacob, Andreas Lanz, Daniel Shapira, and Florian Stahl (2022), "The Research Behind Influencer Marketing," *Impact at JMR* (February 16), <https://www.ama.org/2022/02/16/the-research-behind-influencer-marketing/>.
- Granovetter, Mark S. (1973), "The Strength of Weak Ties," *American Journal of Sociology*, 78 (6), 1360–80.
- Gu, Xian, Xiaoxi Zhang, and P.K. Kannan (2023), "Influencer Mix Strategies in Livestream Commerce: Impact on Product Sales," *Journal of Marketing* (published online October 28), <https://doi.org/10.1177/00222429231213581>.
- Haenlein, Michael and Barak Libai (2013), "Targeting Revenue Leaders for a New Product," *Journal of Marketing*, 77 (3), 65–80.
- Haenlein, Michael and Barak Libai (2017), "Seeding, Referral, and Recommendation: Creating Profitable Word-of-Mouth Programs," *California Management Review*, 59 (2), 68–91.
- Hellinger, Ernst (1909), "Neue Begründung der Theorie Quadratischer Formen von Unendlichvielen Veränderlichen," *Journal für die Reine und Angewandte Mathematik*, 1909 (136), 210–71.
- Hinz, Oliver, Bernd Skiera, Christian Barrot, and Jan U. Becker (2011), "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing*, 75 (6), 55–71.
- Hofstetter, Reto, Andreas Lanz, and Navdeep Sahni (2023), "Contract Design in Influencer Marketing," working paper.
- Hootsuite (2022), "How to Get More Followers on Instagram [5 Simple Steps]," (July 13), <https://lnkd.in/ejm4HEYv>.
- Hughes, Christian, Vanitha Swaminathan, and Gillian Brooks (2019), "Driving Brand Engagement Through Online Social Influencers: An Empirical Investigation of Sponsored Blogging Campaigns," *Journal of Marketing*, 83 (5), 78–96.
- Kumar, Vineet and K. Sudhir (2021), "Can Random Friends Seed More Buzz and Adoption? Leveraging the Friendship Paradox," SSRN (October), <https://doi.org/10.2139/ssrn.3938101>.
- Lanz, Andreas, Jacob Goldenberg, Daniel Shapira, and Florian Stahl (2019), "Climb or Jump: Status-Based Seeding in User-Generated Content Networks," *Journal of Marketing Research*, 56 (3), 361–78.
- Lanz, Andreas, Jacob Goldenberg, Daniel Shapira, and Florian Stahl (2023), "Buying Future Endorsements from Prospective Influencers on User-Generated Content Platforms," *Journal of Marketing Research* (published online September 29), <https://doi.org/10.1177/00222437231207323>.
- Leung, Fine F., Flora F. Gu, Yiwei Li, Jonathan Z. Zhang, and Robert W. Palmatier (2022), "Influencer Marketing Effectiveness," *Journal of Marketing*, 86 (6), 93–115.
- Meyners, Jannik, Christian Barrot, Jan U. Becker, and Jacob Goldenberg (2017), "The Role of Mere Closeness: How Geographic Proximity Affects Social Influence," *Journal of Marketing*, 81 (5), 49–66.
- Newman, Mark (2010), *Networks: An Introduction*. Oxford University Press.
- Pei, Amy and Dina Mayzlin (2022), "Influencing Social Media Influencers Through Affiliation," *Marketing Science*, 41 (3), 593–615.
- Peng, Jing and Christophe Van den Bulte (2024), "Participation vs. Effectiveness in Sponsored Tweet Campaigns: A Quality-Quantity Conundrum," *Management Science* (published online January 29), <https://doi.org/10.1287/mnsc.2019.01897>.
- Perugini, Marco, Marcello Gallucci, Fabio Presaghi, and Anna Paola Ercolani (2003), "The Personal Norm of Reciprocity," *European Journal of Personality*, 17 (4), 251–83.
- Rajkumar, Karthik, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral (2022), "A Causal Test of the Strength of Weak Ties," *Science*, 377 (6612), 1304–10.
- Rapoport, Anatol (1953), "Spread of Information Through a Population with Socio-Structural Bias: I. Assumption of Transitivity," *Bulletin of Mathematical Biophysics*, 15 (4), 523–33.
- Reis, Harry T., Michael R. Maniaci, Peter A. Caprariello, Paul W. Eastwick, and Eli J. Finkel (2011), "Familiarity Does Indeed Promote Attraction in Live Interaction," *Journal of Personality and Social Psychology*, 101 (3), 557–70.

- Romero, Daniel and Jon Kleinberg (2010), "The Directed Closure Process in Hybrid Social-Information Networks, with an Analysis of Link Formation on Twitter," in *Proceedings of the International AAAI Conference on Web and Social Media*. Association for the Advancement of Artificial Intelligence, 138–45.
- Rossi, Federico and Gaia Rubera (2021), "Measuring Competition for Attention in Social Media: National Women's Soccer League Players on Twitter," *Marketing Science*, 40 (6), 1147–68.
- Schmitt, Philipp, Bernd Skiera, and Christophe Van den Bulte (2011), "Referral Programs and Customer Value," *Journal of Marketing*, 75 (12), 46–59.
- SoundCloud (2023), "Promote and Improve Your Stats," (accessed November 24, 2023), <https://lnkd.in/et4wuGyg>.
- Stephen, Andrew T. and Donald R. Lehmann (2016), "How Word-of-Mouth Transmission Encouragement Affects Consumers' Transmission Decisions, Receiver Selection, and Diffusion Speed," *International Journal of Research in Marketing*, 33 (4), 755–66.
- Swaminathan, Vanitha, Karen L. Page, and Zeynep Gürhan-Canli (2007), "'My' Brand or 'Our' Brand: The Effects of Brand Relationship Dimensions and Self-Construal on Brand Evaluations," *Journal of Consumer Research*, 34 (2), 248–59.
- Tajfel, Henri and John C. Turner (1979), "An Integrative Theory of Intergroup Conflict," in *The Social Psychology of Intergroup Relations*, William G. Austin and Stephen Worchel, eds. Brooks-Cole, 33–47.
- Thomson, Matthew (2006), "Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities," *Journal of Marketing*, 70 (3), 104–19.
- Valsesia, Francesca, Davide Proserpio, and Joseph C. Nunes (2020), "The Positive Effect of Not Following Others on Social Media," *Journal of Marketing Research*, 57 (6), 1152–68.
- Van den Bulte, Christophe, Emanuel Bayer, Bernd Skiera, and Philipp Schmitt (2018), "How Customer Referral Programs Turn Social Capital into Economic Capital," *Journal of Marketing Research*, 55 (1), 132–46.
- Van den Bulte, Christophe and Stefan Wuyts (2007), *Social Networks and Marketing*. Marketing Science Institute.
- Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Watts, Duncan J. and Peter Sheridan Dodds (2007), "Influentials, Networks, and Public Opinion Formation," *Journal of Consumer Research*, 34 (4), 441–58.
- Weiler, Michael, Simon Stolz, Andreas Lanz, Christian Schlereth, and Oliver Hinz (2022), "Social Capital Accumulation Through Social Media Networks: Evidence from a Randomized Field Experiment and Individual-Level Panel Data," *Management Information Systems Quarterly*, 46 (2), 771–812.
- Wies, Simone, Alexander Bleier, and Alexander Edeling (2023), "Finding Goldilocks Influencers: How Follower Count Drives Social Media Engagement," *Journal of Marketing*, 87 (3), 383–405.