Methods for Generating Typologies of Non/use

DEVANSH SAXENA, Marquette University, USA
PATRICK SKEBA, Lehigh University, USA
SHION GUHA, Marquette University, USA
ERIC P. S. BAUMER, Lehigh University, USA

Prior studies of technology non-use demonstrate the need for approaches that go beyond a simple binary
distinction between users and non-users. This paper proposes a set of two different methods by which
researchers can identify types of non/use relevant to the particular sociotechnical settings they are studying.
These methods are demonstrated by applying them to survey data about Facebook non/use. The results
demonstrate that the different methods proposed here identify fairly comparable types of non/use. They also
illustrate how the two methods make different trade offs between the granularity of the resulting typology and
the total sample size. The paper also demonstrates how the different typologies resulting from these methods
can be used in predictive modeling, allowing for the two methods to corroborate or disconfirm results from
one another. The discussion considers implications and applications of these methods, both for research on
technology non/use and for studying social computing more broadly.

CCS Concepts: • Human-centered computing → Collaborative and social computing design and
evaluation methods; HCI design and evaluation methods; • Social and professional topics → User character-
istics.

Additional Key Words and Phrases: Methods, non/use, surveys, modeling

ACM Reference Format:
Devansh Saxena, Patrick Skeba, Shion Guha, and Eric P. S. Baumer. 2020. Methods for Generating Typologies
of Non/use. Proc. ACM Hum.-Comput. Interact. 4, CSCW1, Article 27 (May 2020), 26 pages. https://doi.org/10.
1145/3392832

1 INTRODUCTION

A growing body of arguments and evidence make clear the insufficiency of a binary distinction
between “users” and “non-users.” Prior work describes myriad forms of engagement with and
disengagement from technology [6, 8, 10, 23, 24, 56, 66, 90–92, 103].

It is less clear, though, exactly what kind of an alternative to a binary distinction researchers
should employ. Despite the growing body of research cited above, work in this area has yet to
articulate a single, definitive typology of non/use1.

The first two authors contributed equally to this research.

1Following Baumer et al. [9], this paper employs the term “non/use” as a shorthand for “use and non-use.”

Copyright held by the owner/author(s). Publication rights licensed to ACM.
2573-0142/2020/5-ART27 $15.00
https://doi.org/10.1145/3392832
This situation likely arises in part from the fact that the types of relationships (i.e., the subject positions [8]) in question are often specific to particular technologies and social contexts. Reddit is unique in its policy condoning and even encouraging throw-away accounts [66]. Facebook’s deactivation mechanism creates a liminal status [6]. Grindr’s location-based nature grants it varying salience depending on a user’s geographic location [23]. This small handful of examples helps illustrate the difficulty in developing a single, canonical typology to account adequately for the myriad forms of sociotechnical dis/engagements observed in practice.

This lack of means by which to agree upon a typology for forms of non/use can become an impediment to certain types of research. While qualitative research can illuminate the above described complexity, quantitative research that makes use of inferential statistics such as Generalized Linear Models (GLMs), many times based on analysis of survey data, often requires a finite number of categories to converge successfully [72]. These categories must also have prescriptive, unambiguous, mutually exclusive criteria by which each participant can be assigned to exactly one category. Such categories almost necessarily gloss over the variety of experiences and relationships that study participants have with technology. Indeed, in trying to analyze a dataset about Facebook non/use, the authors of this paper found themselves in such a situation, caught between the Scylla of over-simplified classification schemes and the Charybdis of rich but uncodifiable collections of practices. We suspect that other researchers studying non/use likely face similar challenges.

To address these difficulties, this paper contributes a set of two distinct methods for generating typologies of non/use. The first method uses negative binomial curve-fitting to identify the most prevalent types of non/use. The second method generates and assesses different taxonomic trees for assigning a non/use type to each respondent. These methods are demonstrated using survey data related to Facebook non/use. For each method, the paper describes both the process and the rationale for each step therein. The results show how the two methods proposed here generate typologies that are almost directly comparable. These two typologies are then each used to develop separate statistical models for predicting each survey respondent’s non/use type. We demonstrate how alignments between the two models’ results can serve to increase confidence in those specific results, while divergence can introduce a healthy skepticism. We also compare these results with a previous a priori typology from the literature that had not been developed in a data-driven fashion [11]. The results show a strong alignment between that a priori typology and the data-driven typologies generated here, more so for the taxonomic tree typology than for the curve-fitting typology. The discussion addresses how similar approaches could be adapted not only to examining technology non/use in different contexts, but also to much broader aims of generating typologies of human behavior in a variety of domains.

2 RELATED WORK

2.1 Types of Technology Non/use

Numerous different approaches have been taken to classifying different types of technology use and non-use. Much of the earliest work around non-use is framed in terms of a digital divide [103]. Many studies, especially surveys, continue to treat non-use in a binary fashion [1, 3, 54, 89, 95, 97].

Others, however, have offered approaches that go beyond a simple dichotomy [7, 9]. Lenhart and Horrigan [68] argue that the digital divide in terms of internet access should be seen not as a binary but as a spectrum. This spectrum includes, among others, the truly unconnected, who have no possibility for internet access; dropouts and evaders, who could have internet access but choose not to; and home broadband users. While not following the specific types articulated by Lenhart and Horrigan [68], more recent work has similarly considered intensity of use or non-use as a factor [55, 65, 99].
Wyatt [103] combines the dimension of volitionality - willingly choosing to forgo technology - with temporality. Doing so differentiates between resisters, who have never adopted a given technology, and rejecters, who previously used the technology but have ceased doing so. Satchell and Dourish [91] list six different potential types of non-use, from lagging adoption, to disenchantment, to disinterest. This approach emphasizes not only the form but the sociocultural significance of the technology and/or its non-use. Similarly, researchers have started investigating how technology non/use practices vary between different ethnic and socio-economic groups [43], with respect to different IoT devices [42], as well as users’ adoption of non-use practices in response to a company’s values and/or actions [69].

Harmon and Mazmanian [56] articulate four different subject positions of use and non-use, ranging from the multi-tasking master to the out-of-touch luddite. They argue that, in practice, many individuals are in a constant state of tension. Continuously negotiating their engagement with and disengagement from technology causes them to fluctuate among these different subject positions. Many of these prior typologies have appealing qualities. For example, practices such as lagging adoption [91], resistance [103], or the multi-tasking master [56] could conceivably occur with any of a variety of technologies.

However, whether or not each of this practices actually occurs is a different question. Experiences such as “fading away” [23] have been noted only with certain technologies. Similarly, only some work has analyzed cases of individuals reverting to a technology that they had previously ceased using [10, 92].

Furthermore, many systems have unique technical capabilities for enabling non/use, including limited use. Facebook allows users to deactivate, which hides their profile from other users but retains their data on Facebook’s servers for reactivation at any time [38]. Some users attempt to engage with technology only in ways that they find meaningful [70] or employ one set of technologies to control their use of others [71, 83]. Returning to social media after a break is sometimes associated with pruning one’s friend list [10, 13, 92]. Such unfriending involves different steps and has different social meanings among different social media platforms [19, 45, 63, 64].

This variety makes clear two points. First, documented forms of non/use are both numerous and increasing in number with almost every study. Many of those forms are specific to a given sociotechnical setting or context of non/use. It is also unclear how best to compare results across different studies when such varied approaches are used to identifying non/use types.

Second, researchers lack a rigorous means of determining what counts (or perhaps should count) as a distinct type or form of non/use, either in a specific setting or more generally. Put differently, there is no clearly principled way to determine which non/use practices one should expect to observe in any given sociotechnical system. To echo sentiments from the introduction, asking about every previously-observed form of non/use likely creates distinctions that are either too general to be informative or too fine-grained to be meaningful. Thus, rather than create a single typology that could apply to any situation, this paper instead offers methods by which researchers can develop a typology specific to their research setting.

2.2 Methods for Typology Development

Various related bodies of work in HCI, CSCW, and social computing have had differing approaches to thinking about typology development. Some of this variety results from drawing on a varied set of “parent” disciplines, such as psychology, sociology, communication, computer science, information science etc.

In general, typology development can refer to two distinct scientific activities. First, in a broad sense, it may refer to theory building that centers around classifying phenomena in different ways. This is usually (but not entirely) qualitative and draws from existing literature to construct a
grounded notion of how an aspect of the world works or is explained via sorting and categorization. For instance, within HCI/CSCW/social computing, typologies drawn from largely qualitative research practices such as focus groups, interviews, ethnographies, and exploratory cluster analysis have been developed in a wide variety of domains such as blockchain applications [37], universal design principles for software applications [40], online engagement [101], and social networking site users [22]. A meta analysis of media-user typology research confirms the above suite of methods for developing typologies as common practices in this broad discipline [21]. A common critique of these approaches is that they are hard to generalize between and across various domains or context withing HCI/CSCW/social computing [78].

Second, in a more specific sense, typology development may refer to the methodological development of theoretical categories into classification schemes and is usually facilitated via quantitative methods. There exists a wide body of literature within statistics and machine learning for optimizing the number of classes and classification rules. General methods include multi-class support vector machines [100], various ensemble learning practices [31], and of course, the more recent deep learning methods [28]. However, one common critique with the above range of methods is that they engage less with theory and are hard to interpret [32], especially when developing typologies or classes purely from a mathematical optimization perspective.

This is not to say that efforts have not been made within HCI/CSCW/social computing to reconcile two different methodological approaches to the same problem [12, 78]. For instance, very recently, Latent Profile Analysis (LPA) and Latent Class Analysis (LCA) have been proposed to bridge this gap [80]. However, this approach depends on a thorough knowledge of mixture models which are technically more complex and can be harder to interpret especially when comparing the classes between two typologies [27]. We examined other scientific domains such as environmental engineering and energy forecasting that have made good use of curve-fitting [44] and computational taxonomic methods [61] to bridge the qualitative and quantitative method gap described above. We adapt these methods in our study to develop a generalizable and more accessible method (that offers higher transparency and interpretability) for HCI/CSCW/social computing researchers studying complex theoretical constructs.

3 CASE STUDY: DEVELOPING FACEBOOK NON/USE TYPOLOGIES

The above literature describes varied forms of Facebook (and, more broadly, social technology) non/use. These forms were used to craft the following Yes/No questions as survey items. Following Wyatt [103], each respondent can be labeled as engaging or not engaging in each of these activities. Parenthetical text, which was not included in the survey, provides additional details relating each item to prior studies.

- I currently have an active Facebook account.
- I have more than one Facebook account. (Multiple accounts are sometimes used to manage or limit one’s own Facebook use and/or others’ access to one’s online profile [20, 74, 75]).
- At some time, I have deactivated my Facebook account. [6, 20, 92].
- At some time, I have permanently deleted my Facebook account. [6, 85].
- At some time, I have voluntarily taken a break from Facebook for a week or more. [6, 10, 87, 92]
- At some time, I have used software to limit my Facebook usage. [83, 92]
- At some time, I have deleted the Facebook app from my phone. [10]
- I have never had a Facebook account. [33]
- It is my own choice whether or not I use Facebook. (Prior work has documented where people, against their will, either were prevented from using Facebook or felt forced to have an account [6, 13, 91, 103–105])

Proc. ACM Hum.-Comput. Interact., Vol. 4, No. CSCW1, Article 27. Publication date: May 2020.
To be sure, this typology is non-exhaustive. For example, it does not include politically-motivated technology abstention [85], gradual “fading away” of a technology [23], constant feelings of tension and instability [56], or asking a friend to change one’s password [6]. Instead, it focuses on the most common forms of Facebook non/use documented in prior work that could conceivably be captured with Yes/No questions. The correlation matrix of these questions can be found in the appendix.

A survey design was sought that would encompass as many of these different practices as possible. However, prior literature provided little a priori expectation about how these individual practices might be organized into a coherent typology. For example, should deactivating one’s account be seen as a special case of taking a break from Facebook? Should it be assumed that all respondents who deleted their account also deleted the Facebook app from their phone? What if a respondent with multiple Facebook accounts deactivated one account, deleted another, and left a third account active?

Instead of making such decisions at the outset, each respondent answered every one of the Yes/No questions above. Not only does this approach obviate the need to make assumptions about relationships among these practices, it also extends Wyatt [103]. She used two dimensions, temporality (used technology before or did not) and volitionality (non-use is a willful choice or not), resulting in a 2-dimensional grid with a total of $2^2 = 4$ types of non/use. We extend this idea to generate a N-dimensional grid with a total of $2^T$ types, where $T$ is the number of Yes/No questions included. Researchers studying different data sets or working in other domains could generate their own list of Yes/No questions based on relevant prior literature and expectations, a point considered further below in the Discussion.

The responses to these Yes/No questions were used to generate typologies from the bottom-up. There is little reason to expect that at least one participant will occupy each and every one of the total $2^T$ (in this case, $2^9 = 512$) types of non/use. Furthermore, some types may never occur; we would expect it impossible that a single participant both never had a Facebook account and currently has an active Facebook account. Thus, we propose a data-driven approach to identify the most perspicacious typology possible, as described below in Section 4.

Using two different methodological lenses, we derived two typologies based on participants’ responses to the seven typological survey questions (both typology generation methods resulted in dropping two of the typological survey questions, as described below). For each of these typologies, we then trained a multinomial logistic regression model to predict non/use type – one separate model for each typology – using features extracted from the psychometric scales included in the survey. The goal was to evaluate whether these two different methods would converge towards a single, consistent non/use typology or generate divergent, non-comparable typologies.

3.1 Data

Data were collected as part of a larger survey about Facebook non/use. Participants were recruited via Qualtrics, whose recruitment and sampling procedure is outlined on their website (https://www.qualtrics.com/online-sample/). Qualtrics’ staff assembled a web panel of participants with a demographic composition resembling that of general internet users with respect to age, education, and ethnicity [82]. Respondents were screened at the beginning of the survey using these criteria. For example, once the survey received 88 respondents age 25-34 (i.e., 17.7% of our target sample size of 500 respondents), subsequent respondents in the age 25-34 did not pass the age criterion. Respondents who did not pass any of the demographic screening criteria were excluded. Ultimately, the web panel included 516 participants (with 1028 potential respondents screened for not passing demographic criteria), for which we paid $2,500.

The survey also included a variety of other items that might relate to Facebook non/use, mostly drawn from prior related work. The Facebook Inventory (FBI) [36] was used to assess intensity
of Facebook usage [cf. 55]. A prior uses and gratifications analysis found five distinct factors in why people use Facebook [67]: to combat survey fatigue, the present survey included only the two highest-loading questions for each of these factors. The six-item version of the Bergen Facebook Addiction Scale (BFAS) [2] was used to assess limited impulse control [48, 49]. A ten-item scale [46] assessed personality [29], which prior work has found to differ in binary comparisons of users and non-users [89, 95]. Since prior work has also found an individual’s subjective feelings of control to be an important factor [13, 56, 92, 102], the survey included a validated scale for individual sense of agency [96]. Since multiple prior studies have linked privacy and non/use [1, 6, 87, 95, 97], the survey included several questions about disclosure of potentially private information [58, 98]. The survey included several self-report questions about respondents’ social networks; in the interest of simplicity, these are omitted from the present analysis. Several questions had optional free-text fields for participants to further elaborate a response; for instance, participants who indicated that they had multiple accounts were prompted to explain why they did so. The free-text responses were not used in this quantitative analysis, except to check whether participants were interpreting the questions in the manner we intended. Finally, demographics included gender (male, female, other), household income, zip code, marital status, and political views (7-point Likert from Very Liberal to Very Conservative).

4 TWO METHODS OF TYPOLOGY GENERATION

Due to the absence of definitive methodologies for generating typologies of non/use, we opted for a data-driven approach. We considered two methods, based on different statistical techniques, and explain here the tradeoffs that each method entails. In the subsequent section, we present the results of an experiment in which we train a multinomial logistic regression model to predict classes in each of the typologies based on the survey data. This experiment helps to illustrate the value of each typology and encourages us to suggest that future work consider multiple non/use typologies. Such an approach should prove especially useful when, as in the case of Facebook, no definitive classes of non/use exist. We also provide a blueprint (illustrated in Figure 5) for this typology generation process which will allow researchers working in other domains to replicate our methodology for generating classes.

4.1 Method 1: Curve-fitting solution

Broadly speaking, the negative binomial curve fitting approach identifies every unique combination of Yes’s and No’s across all the typology questions. We dropped two of the nine typological questions because no (or very few) users responded affirmatively to these two questions. “I have never had a Facebook account” represents zero users and “It is my own choice whether or not to use Facebook” represents less than 2% of the users.

Each unique combination of Yes’s and No’s to the remaining seven questions becomes a type of non/use. Our initial analysis of the 7 typological questions revealed 50 unique user-behaviors. These questions were borrowed from prior literature as being significant in identifying types of non/use. However, we sought to avoid any a priori assumptions, either about which of these practices do and do not co-occur, or about their relative prevalence. Therefore, we use this data-driven approach to figure out the most prevalent types of non/use. This approach then attempts to identify the optimal number of types to include, balancing the number of respondents who fit into one of those types (more is preferred) against the total number of types (fewer is preferred). For this step, we focused

---

2Since there is some debate over the suitability of “addiction” to describe potentially problematic social media usage [30, 51, 52, 84], BFAS is used here as a measure of a respondent’s ability to control their own usage of Facebook, as well as negative consequences that may arise from high levels of usage.
Table 1. Typology generated by curve-fitting method. The "+" sign indicates that the users had engaged in the specified non/use behaviors at some point in the past. The sign does not establish an ordinal or temporal relationship between the non/use behaviors.

| Type | Users |
|------|-------|
| FB   | 127   |
| FB + takenBreak | 99 |
| FB + takenBreak + deactivated + deletedApp | 44 |
| FB + takenBreak + deletedApp | 34 |
| FB + takenBreak + deactivated | 21 |
| FB + takenBreak + deactivated + deletedApp + deleted | 21 |
| FB + takenBreak + deactivated + deletedApp + FBmorethan1 | 14 |
| FB + FBmorethan1 | 14 |
| takenBreak + deactivated + deletedApp + deleted | 10 |
| FB + deactivated | 10 |
| FB + takenBreak + FBmorethan1 | 9 |
| FB + takenBreak + deactivated + deletedApp + usedSoftwareToLimit | 9 |
| FB + takenBreak + deactivated + deletedApp + deleted + usedSoftwareToLimit | 9 |
| deactivated + takenBreak + deletedApp | 7 |
| FB + deletedApp | 7 |

FB: I currently have a Facebook account
FBmorethan1: I have more than one Facebook account
deactivated: At some point, I have deactivated my Facebook account
deleted: At some point, I have permanently deleted my Facebook account
takenBreak: At some point, I have voluntarily taken a break from Facebook for a week or more
usedSoftwareToLimit: At some time, I have used software to limit my Facebook usage
deletedApp: At some point, I have deleted the Facebook app from my phone

on the top fifteen most common non/use types, as depicted in Table 1. After the top fifteen non/use types, very few participants (less than 6) belonged to the classes that fell in the long tail of the distribution.

Next, we construct a mathematical curve that best fits the relationship between the number of respondents of each non/use type and the frequency of each non/use type in the data set (see Figure 1). A negative binomial distribution is the best fit theoretical distribution for this relationship. Figure 1 depicts the fitted mathematical function represented by the following equation where $x$ represents each non/use type (ordered by frequency in the data set) and $y$ is the number of users corresponding to that non/use type.

$$y = 10.635 + 150.740e^{-((x+0.604)/3.271)^2}$$

In accordance with the Pareto principle [18], we focus on the non/use behavior of 80% of the respondents. Our research design uses Pareto’s 80/20 rule but it is not a compulsion, and other studies may deviate from this ratio based on the number of classes present on the left side of the equation. Generalized Linear Models are built using maximum likelihood estimation and their performance deteriorates as the number of classes on the left side of the equation increases [72]. Therefore, it is imperative to limit the number of classes on the left side of the regression model, i.e., the number of distinct types of non/use emerging from the typology. Figure 1 shows that the top six non/use types can be used to describe the non/use behavior of 80% of the respondents.
Fig. 1. Curve-fitting solution describing non-use behavior

C1: FB
C2: FB + takenBreak
C3: FB + takenBreak + deactivated + deletedApp
C4: FB + takenBreak + deletedApp
C5: FB + takenBreak + deactivated
C6: FB + takenBreak + deactivated + deletedApp + deleted

Note: The "+" sign indicates that the users had engaged in the specified non/use behaviors at some point in the past. The sign does not establish an ordinal or temporal relationship between the non/use behaviors.

4.2 Method 2: Taxonomic-Tree solution

As a complement to the above curve-fitting solution for determining typologies, one could also use a taxonomic tree to select a subset of typology questions and place all respondents into exactly one category. The basic process is to choose a question on which to divide the population into the 'Yes' and 'No' groups, and then iteratively subdivide those groups on new questions until either all nine questions are used or one of the subdivisions has too few members, according to some criteria. We chose 40, or approximately 8% of the 514 total respondents, as the minimum and stopped the iterative splitting if it resulted in groups with less than this number of members.

Choosing a Taxonomy-based Typology – The process of selecting a taxonomic tree to define the typologies is nontrivial. Enumeration of all possible trees is known to be NP hard and could become infeasible as the number of questions in consideration grows, necessitating the use of more sophisticated algorithms like the one described in [88]. In our case, a naive approach was sufficient, but this issue of scalability should be noted. By requiring a minimum of 40 members in each group we restricted the number of possible trees, resulting in a total of 329 candidates.

The second challenge was deciding which of the 329 candidate trees to use in defining the typologies. To our knowledge, the problem of selecting the "best" taxonomic trees for use as an outcome variable in a regression model is an unexplored area of research. Thus, we devised evaluation criteria that fit our dataset. First, we excluded all trees which contained a leaf with fewer
than 51 members (or 10% of the total) on the bases that smaller groups would create challenges for the regression models and risk over-fitting. This significantly reduced the number of candidates from 329 to 59.

Next, we applied an upper bound on the size of each group of 172, or one third of the total number of respondents. This number was chosen to ensure that we had at least two splits in the tree, i.e., at least three groups. This step excludes typologies based on a single question, such as simply comparing those respondents who have deleted their account vs. those who have not. Application of this criterion resulted in 20 candidate trees.

Half of the remaining 20 trees included splits based on having multiple accounts. Upon closer inspection of those respondents who answered “yes”, we saw significant disagreements in how the question was interpreted based on the associated free-text fields. Some respondents considered themselves to have multiple accounts because they deleted theirs and created a new one, others noted that they were forced to make a new account because they could not log in to their prior account. Due to the ambiguity of the question as it was worded, we chose to exclude trees which split the data based on maintenance of multiple accounts.

After removing trees containing the multiple accounts question, we were left with 10 candidates. Finally, we sorted these remaining trees by the variance of the leaves, in order to generate a taxonomy with the most equally sized groups. The final selected typology is depicted in Figure 2.

### 4.3 Comparing the Two Methods

The curve-fitting method is an entirely data-driven process that allows us to recognize and model the most prevalent types of non/use while ignoring the more rare types that end up in the long tail of the distribution. Therefore, if researchers were interested in especially studying one of these rare classes, the curve-fitting method would perform poorly because the learning algorithm may fail to identify patterns in the rare classes because of over-representation of the more prevalent classes. For instance, in our survey less than 2% of the participants responded affirmatively to the question, "It is my choice whether or not to use Facebook". So, researchers who want to study volitionality (non-use is a willful choice or not) would want to use methods (for e.g., taxonomic trees) that allow them to generate typologies that focus on rare classes. Moreover, the curve-fitting method allows us to study the prevalence of non/use behaviors as a general theoretical construct, however, it does not allow us to investigate specific non/use behaviors.

On the other hand, the taxonomic tree method is a data-driven process that allows for qualitative interventions based on the research criteria. In our case, we devised evaluation criteria based on the intensity of non/use behaviors. As mentioned earlier, enumeration of all possible trees in an NP
hard problem, therefore, researchers must theoretically design criteria that allows them to generate specific trees that focuses on a specific phenomena. Making these decisions inadvertently means that we will be losing some information since we decided to focus on certain trees that offer the most information about a phenomena. For instance, in our case we had to choose from a subset of the non/use questions that offer the most information with respect to the intensity of non/use behaviors.

These two methods used in conjunction may offer pathways for generating typologies for studying gender identity and sexual orientation. As non-binary genders are becoming increasingly prevalent, it becomes important to be able to study both the prevalence of different genders (curve-fitting) as well as gender variation within non-binary genders (taxonomic trees) without othering or erasing non-binary participants who might associate with types with smaller class sizes [39, 57, 94].

---

Proc. ACM Hum.-Comput. Interact., Vol. 4, No. CSCW1, Article 27. Publication date: May 2020.
Methods for Generating Typologies of Non/use

4.4 Comparing the Two Generation Methods and Typologies

The most significant consideration when comparing the two methods is the tradeoff between the size of the dataset used for training and the fineness of the groups derived in each typology. The curve-fitting method uses all seven questions to create a class for each unique set of responses, but excludes datapoints in order to discount the long tail and maintain large enough class sizes to make classification feasible. On the other hand, the taxonomic method preserves the number of datapoints, but makes less fine distinctions between classes by only using a subset of the questions. For all classes in the curve-fitting typology, all seven questions are used and the average class size is 58 ($n = 346$), while the taxonomic-tree typology uses a maximum of four questions (T1 and T2) and a minimum of one (T5) with an average class size of 103 ($n = 514$).

In terms of the resultant typologies, a theoretical mapping exists between most of the types. For instance, as depicted in Figure 5, the curve-fitting type C6 maps onto the taxonomic-tree type T5 because participants who engage in the more intense forms of non/use are also more likely to have engaged in the less intense form of non/use. Figure 4 illustrates the non/use behaviors of the 74 participants in our study who had permanently deleted their Facebook account (taxonomic type T4). The majority of these participants engaged in other types of non/use, as well. Figure 5 also illustrates the differences in class size and overall number of datapoints, and suggests that the two methods of deriving typologies have in fact converged to similar non/use types. Each taxonomic type has a corresponding curve-fitting type, the latter of which makes up a subset of the former.

5 TYPOLOGY COMPARISON ON SURVEY DATA CLASSIFICATION TASK

Using the typologies described above, we conducted two nearly-identical analyses. Similar to Baumer [5], both analyses used multinomial logistic regression, where the predictors were attributes of each respondent and the outcome variable was the respondent’s non/use type. The only difference was whether each respondent’s non/use type (i.e., the outcome variable) was assigned based on the curve-fitting typology or on the taxonomic-tree typology.

This analysis involved a series of steps. First, predictors were extracted from the survey data by performing factor analysis to reduce the dimension of the psychometric scales and including demographic information. Second, model selection was then used to limit predictors to only those which significantly improve our information about the outcome variable. Finally, to compare the two typology generation methods, we identified points of alignment and of divergence, both in terms of which predictors are maintained after model selection, and in terms of the direction and

![Fig. 4. Most participants who permanently deleted their Facebook account (taxonomic type T4) engaged in at least one other non/use behavior (deactivating their account, deleting the app from their mobile device, or taking a break).](image-url)
magnitude of those predictors’ effects. The remainder of this section describes each of these steps in detail.
5.1 Feature Extraction from Psychometric Scales

With either the curve-fitting typology or the taxonomic-tree typology serving as the left side of the regression model (i.e., the outcome variable), we used the five psychometric scales included in the survey to extract features for the right side. For four of the scales, exploratory factor analysis (EFA) was used to identify underlying structures and restrict the number of predictors in the final model. We did not perform factor analysis on the fifth scale, the 10-item Big 5 Personality Scale, choosing instead to use its intended scoring procedure to create a predictor for each of the five personality dimensions. This decision was made after observing that the results from our own factor analysis on the Big 5 items did not match the standard interpretation, which has been validated on large datasets and across different demographics and cultures [15, 59, 60].

All exploratory factor analyses were performed on the entire 514 member dataset in order to extract the same features with which to train a classification model for both the curve-fitting and taxonomic typologies. Predictors were constructed for each psychometric scale using an average of all questions loading on a given factor, weighted according to the factor loadings. Cronbach’s alphas are provided to show internal consistency of factors, and the full factor structures can be found in the appendix.

5.1.1 Bergens Facebook Addiction Scale (BFAS). Andreassen et al in [2] found six underlying factors using the original 18-question version of the scale which correspond to six dimensions of addictive behavior: salience (the behavior dominates one’s thinking and behavior), tolerance (more of a behavior/activity is needed to achieve desired effects), mood modification (activity affects/improves how one is feeling), withdrawal (unpleasant feelings when activity is halted or reduced), conflict (activity causes interpersonal conflict or interferes with other activities), and relapse (repeated reversions to past behaviors after limiting or abstaining) [2, 50]. Our analysis suggested three underlying factors for this dataset: a combination of Salience, Tolerance, and Mood ($\alpha = .83$), a combination of Withdrawal and Conflict ($\alpha = .84$), and Relapse. 73% of the cumulative variance is explained by these three dominant factors.

5.1.2 Facebook Intensity Scale (FBI). Factor analysis was performed on the six FBI questions related to how a person feels about their Facebook use. The remaining two questions on number of Facebook friends and time spent on Facebook were not included in the factor analysis and were treated as individual predictors in the regression models. We found that three factors, which we call Connectedness ($\alpha = .86$), Daily Routine ($\alpha = .93$) and Pride, explains 81% of the cumulative variance in the full dataset.

5.1.3 Sense of Agency (SoA). EFA on the Sense of Agency survey questions suggested two factors, Sense of Positive Agency (SoPA) ($\alpha = .79$) and Sense of Negative Agency (SoNA) ($\alpha = .84$), which align with the factors identified in [96]. 43% of the cumulative variance is explained by these two dominant factors and did not significantly improve with more factors.

5.1.4 Leiner et al.’s Uses and Gratifications. Table 9 depicts the factor analysis conducted using questions from Leiner et al. [67]. We see Social integration ($\alpha = .91$), Affective gratification ($\alpha = .91$), Personal integration ($\alpha = .84$) and Escape ($\alpha = .81$) emerge as the dominant factors. These factors were identified based on the major items referenced in Leiner et al. [67]. 68% of the cumulative variance is explained by these four dominant factors.
5.2 Predicting Curve-Fitting and Taxonomic-Tree Typologies with Multinomial Logistic Regression

Finally, we used the predictors derived from EFA on the psychometric scales and demographic items to train two multinomial logistic regression models, one that predicts non/use type according to both the curve-fitting typology (C1-C6), and one that uses the taxonomic-tree typology (T1-T5). In order to avoid overfitting and identify only predictors which significantly increase our knowledge about the outcome variables, we performed model selection to reduce the number of predictors.

5.2.1 Model Selection: Backward Step-wise Regression. We conducted step-wise model selections based on AIC for both the curve-fitting and taxonomic-tree typologies. We used a backward elimination process where we start with all predictors depicted in Table 2 and iteratively eliminated predictors whose removal caused the most statistically insignificant deterioration of the model fit. This process helped us identify the most salient predictors for the multi-logit regression models that will be used to classify users into non/use behaviors. Table 3 depicts our final predictors for the curve-fitting and taxonomic-tree models along with the odds-ratios for all predictors corresponding to each typology.

5.2.2 Model Alignments and Divergences. As expected, the reduced models align on some predictors while diverging on others. In terms of alignment, Relapse and Age are maintained in both models and exhibit similar effects, both in terms of magnitude and in terms of direction. For example, in both models, older respondents were less likely to be anything other than a current user. In terms of divergence, Daily Routine and Affective Gratification are dominant predictors in both models, but their odds-ratios suggest that they have opposite impacts. With the curve-fitting typology, increases in either of these variables predicted increased probability of a respondent being something other than a current user. In contrast, with the taxonomic-tree typology, increases in either of these variables predicted decreased probability of a respondent being something other than a current user. Furthermore, five predictors are retained only with the curve-fitting model, and four predictors are retained only for the taxonomic-tree model. The following discussion interprets these alignments and divergences, situating them in related work on social media use and non-use.

Table 2. Psychometric scale predictors. Each column is based on the factor structure extracted in Section 4.2.

| Y = | BFAS + | Leiner + | Big5 + | Agency + | FBI |
|-----|--------|----------|--------|----------|-----|
| Curve-fitting Typology (C1-C6) | SalTolMood | Social integration | Extraversion | SoPA | Connectedness |
| or | WithdrConfl | Personal integration | Conscientiousness | SoNA | Daily Routine |
| or | Relapse | Escape | Openness | | Pride |
| Taxonomic-tree Typology (T1-T5) | | Affective gratification | Neuroticism | FBI-Time | | Agreeableness | FBI-Friend |

Proc. ACM Hum.-Comput. Interact., Vol. 4, No. CSCW1, Article 27. Publication date: May 2020.
6 DISCUSSION AND CONCLUSION

As this is a methodological paper, the discussion focuses primarily on how the methods proposed here can generate useful insights. The contribution is not necessarily the specific insights per se into Facebook non/use that arise from the classification models. Rather, the contribution is a demonstration of how these two typology generation methods can be leveraged in combination to produce informative findings.

Thus, this discussion first considers implications of these typology methods, and how investigating points of alignment and of divergence in the models can lead to valuable observations. Second, we note important limitations that other researchers should bear in mind when using the methods proposed here. Third, the discussion demonstrates that these methods can in fact produce valuable insights by describing the implications of this work for research on technology non/use. Finally, it considers broader implications for the use of quantitative methods to study technology and social behavior.

6.1 Demonstrating Alignment and Divergence Across Models

Rather than viewing the curve-fitting and taxonomic tree approaches above as competitors between which one must choose, we instead present a holistic methodological approach that compares results across both typology generation approaches (see Figure 5). This section draws on Table 3, which depicts the odds-ratios for the reduced multi-logit models. Recall that the reference categories, C1 and T1, correspond to current Facebook users who have not engaged in any of the non/use practices included in the survey. Therefore, odds-ratios represent the relative probability of a respondent belonging to one of the non/use categories (deletion, deactivation, deleting the app, etc.), compared to this baseline. Examining these odds ratios identifies points of alignment between the two models, reinforcing our confidence in the results. It also draws attention to points of divergence, which can introduce healthy skepticism about single results.

For example, Relapse (returning to old habits after a period of abstinence or reduced Facebook use) shows clear alignment across the two models. The corresponding odds-ratios are greater than one and of a similar magnitude across all classes in both typologies. That is, higher reported Relapse predicts a higher probability of a respondent engaging in some form of non-use behavior. The
similarities in effects and significance (curve-fitting: \( p < .05 \), taxonomic-tree: \( p < .001 \)) increase our confidence in the predictive power of this feature. These findings also align with prior work on taking breaks from social media and returning before one had intended [10, 87, 92]. Furthermore, Age was maintained for both taxonomies with very high significance (\( p < .001 \)); odds-ratios suggest that age is negatively correlated with non-use behaviors (non-adopters were not present in this particular dataset), echoing Baumer [5]. These external resonances lend further credence to our claim that alignment between the two models should increase researchers’ confidence in the results.

On the other hand, there are at least two ways that the results from the two typology generation methods suggested here may diverge from one another. First, model selection may cause certain predictors to be maintained in a model with one typology, but not with the other. Sense of Positive Agency, for example, is maintained with high significance (\( p < .01 \)) only in the taxonomic-tree model, while Pride, FBI-Friend, and Escape are only preserved for the curve-fitting taxonomy. Inclusion of the two BFAS factors besides Relapse slightly reduces AIC in the taxonomic-tree model, but low significance scores cast doubt on their importance. Second, the same predictor across the two different models may have odds-ratios in opposite directions. Daily Routine and Affective Gratification (described above) both exhibit this phenomenon. That said, differences in significance scores may lend more credence to the effects suggested by one model over those suggested by the other. For example, although Daily Routine is maintained in both models, its effect is statistically significant only for the model based on the taxonomic tree typology.

Divergence can be interpreted in a number of different ways. First, these predictors may not actually be significant. That is, their inclusion in the final model may occur only due to statistical chance. Second, the divergences might be resolved with a larger dataset, reducing our uncertainty about the relevance of certain predictors. Finally, these divergences may suggest that the typologies themselves are not as similar as they initially appear. Consider, for example, the relationship between T4 in the taxonomic tree, and C3 and C5 in the curve-fitting typology (Figure 5). Both curve fitting classes (C3 and C5) are strict subsets of the taxonomic tree class (T4). However, C3 distinguishes itself from C5 by requiring that the respondent also deleted the Facebook app from their phone. T4 makes no such distinction. The results show that, when including this distinction within the typology, odds ratios for certain predictors (e.g., Affective Gratification, Escape) have opposite directions. Thus, the results question whether a respondent deleting the Facebook app from their phone should be used as a decision criterion in classifying their non/use type. More significantly, perhaps the two sets of classes arising from the two different typologies are not quite as similar as depicted in (Figure 5). Other researchers using the methods proposed here can similarly identify potentially problematic classes in the typologies generated from their data.

Overall, identifying points of divergence in the models draws our attention to important issues that may have been missed if only one typology were used. Furthermore, doing so raises many questions that can be used to develop more robust typologies of non/use in future work.

At a higher level, such comparisons underscore the strength of the methods proposed here. To reiterate, researchers should not adopt only one of either the curve-fitting or the taxonomic approaches. Rather, they should generate typologies using both these approaches, which enables direct comparisons of the typologies themselves. Furthermore, conducting subsequent analyses, such as the predictive modeling presented here, can provide stronger evidence in support of conclusions drawn from the analysis. To wit, \( p \)-values derived from singular studies are noisy, and we hypothesize that adopting a dual methodological approach would yield results that provide more stable outcomes with higher effect sizes [62]. Issues concerning statistical transparency have received increasing attention in HCI and sociotechnical research [16, 26]. This work speaks to these issues and offers an example of methodologies that provide greater confidence in our results. We expand on this in a subsequent subsection.
### 6.2 Comparison against a third set of non/use classes

We also wanted to see how our typologies compared to other non/use typologies published in recent literature. Baumer et al. [11] presented a non/use typology that compares three different forms of non/use with respect to current Facebook users. The three classes discussed in the paper are: (1) current active users; (2) people who deactivated their account; and (3) people who have considered deactivating their account but have not actually done so. We term these B1, B2, and B3, respectively.

A direct comparison with the Baumer et al. [11] typology was not possible, since our study did not incorporate a non/use question that focused on “considered deactivation.” Thus, we compare type B3 with the “taking a break” question in our data, which is used in types C2 and T2. Again, this does not provide a direct comparison, since actually taking a break involves action, while considering deactivation does not. However, all of B3, C2, and T2 represent the least intense forms of non/use within their respective typologies. Furthermore, in each case, the user continues to maintain an active channel to their Facebook account. We followed the methodology specified by Baumer et al. [11] to assign non/use types (B1, B2, or B3) to our participants. We then trained a multinomial logistic regression model followed by a backward step-wise regression model. Table 4 depicts the significant predictors and their odds-ratios determined by the reduced multi-logit model.

Similar to the curve-fitting and taxonomic-tree typologies, Relapse and Age are significant predictors for Baumer et al.’s typology, with odds-ratios exhibiting similarity in both magnitude and direction. Daily Routine and Affective Gratification are also significant predictors, although the magnitude and direction of their odds-ratios aligns with the taxonomic-tree typology and not with the curve-fitting typology. These statistical differences point to the subtle differences that exist between the curve-fitting and taxonomic-tree typologies as stated previously. More importantly, such comparisons allow us (and other researchers) to iterate back and revise the types within each typology to ensure a more robust alignment.

### 6.3 Limitations

First and foremost, this paper has only examined a single data set, and that data set only pertained to non/use of Facebook. The methods presented here may be more (or perhaps less) difficult with other social media. Moreover, this study only focuses on a subset of non/use types that could conceivably be captured with Yes/No questions (i.e., discrete non/use actions). Furthermore, by focusing on a single social media platform, we were unable to investigate the broader ecology of

---

**Table 4.** Odds ratios for reduced multi-logit model with Baumer et al.’s non/use types [11] as the classification target. Backwards step-wise regression reduced the model to four predictors. Asterisks represent significance level of likelihood-ratio tests on the corresponding predictor ($0.05, 0.01, 0.001$)

| Scale | Predictor                  | Odds-Ratios (Baumer et al. non/use types) |
|-------|---------------------------|------------------------------------------|
|       | B2                        | B3                                       |
| BFAS  | Relapse                   | 1.37***                                  |
| FBI   | Daily Routine             | 0.73**                                   |
| Leiner| Affective Gratif.         | 0.89*                                    |
| Demog | Age                       | 0.97***                                  |

*B1: I currently have a Facebook account (reference category)*

*B2: At some point, I have considered deactivating my Facebook account*

*B3: At some point, I have deactivated my Facebook account*
non/use where users have a choice of several social media platforms [47]. That said, the relative consistency of the results generated here suggest that the process is fairly robust for generating and investigating non/use typologies. We hope that future work will both apply and adapt the methods offered here.

Second, this method focuses purely on survey data and only uses binary Yes/No questions. As described above, this choice was informed by foundational work on different forms of non-use [103]. Future work could expand these methods to include questions with ordinal or continuous responses. Similarly, the methods could be expanded to include data from usage logs [34]3. Again, this paper lays the groundwork upon which future researchers can pursue such directions.

Third, generating novel typologies may significantly limit the ability to compare results across studies. If every data set is used to generate a one-off typology, it becomes difficult to compare and contrast results using such diverse typologies. As an alternative, it might help to consider the methods proposed here as analogous to the construction of validated scales [35, 41]. Rather than generate an entirely new typology for every data set, researchers could instead follow an iterative process of development and validation across multiple studies. Such processes occur in numerous domains for the development of standardized, validated scales, from personality [46, 76, 77] to privacy [25, 73, 93]. Similarly, different researchers across multiple studies should test, question, and revise typologies generated in the manner described in this paper.

Finally, there are two general limitations of this type of survey study that deserve mention. First, we can only capture participants’ opinions at one instant in time and rely on their recall of past behaviors and characteristics of their Facebook use. This limits our ability to understand temporality, or changes in non/use over time. Other studies have attempted to address such aspects using, e.g., repeated sampling of the same population [14] or more longitudinal diary studies [13]. Future work might benefit from combining those techniques with the typology generation methods offered here to examine how typologies themselves change and evolve over time. Second, survey non-respondents may potentially bias our results away from people who are simply less likely to engage in online interaction. This issue similarly occurs in other work targeting individuals who are less directly engaged with technology, such as blog readers [4] or lurkers [79, 86]. We attempted to mitigate this issue by recruiting internet users in a demographically representative sample, rather than Facebook users per se. However, this is likely a fundamental issue with survey research, and other methods would be needed to address it more fully.

6.4 Implications for Technology Non/use Research

This analysis offers multiple specific implications for research on technology non/use. First, at least in the case of Facebook, it might not make sense to consider volitionality as a factor for identifying types of non/use. As demonstrated in the analysis above, responses to the question (about whether respondents felt they had a choice in whether or not to use Facebook) were ultimately not informative. Initially, this result seems to run contrary to prior work arguing for the analytic importance of an individual’s control (or sense of control) over their own technology use [10, 53, 91, 102, 103, 105]. However, these results do not suggest that volitionality, or sense of agency, is unimportant. Rather, they suggest that this factor might be less informative for differentiating among various types of Facebook non/use and more informative as a predictor for an individual respondent’s non/use type. Future work will need to examine the utility of this factor for generating typologies in other sociotechnical contexts.

3In the case of Facebook or other private social media companies, only a limited number of researchers will have access to the details of such log data.
Second, the taxonomic-tree approach suggest some degree of hierarchicality among Facebook non/use types. The most suitable taxonomic-tree prioritizes questions about specific practices in an order that seems to go from most to least intense forms of non-use: from account deletion (permanent and irrevocable), to account deactivation (temporary), to deleting the mobile app (limiting only one access channel), to taking a break (using no technical mechanism).

Such a hierarchy makes intuitive sense and could have been postulated in advance, based on theory and prior work. However, numerous varied hierarchies could be similarly postulated, with little means of choosing from among them. For example, the types of non/use could have been organized in the opposite direction, i.e., from least to most intense forms of non-use. Doing so, however, would have resulted in severe class imbalances, which would have made predictive modeling difficult. Thus, the choice to avoid an *a priori* hierarchy among different non/use practices allowed our data-driven approach to reveal a latent hierarchy with a meaningful interpretation. Future work should consider whether similarly interpretable typologies emerge from other data sets.

Finally, classification systems embody numerous assumptions, value commitments, power dynamics, etc. [17]. The typologies that result from the methods presented in this paper similarly embody commitments and assumptions. Most immediately, the researchers constructing the original Yes/No questions on which the typologies are based have significant influence in what counts as an important factor. More subtly, both the curve-fitting approach and the taxonomic-tree approach favor typologies with fewer types and more respondents in each type. Doing so necessarily marginalizes minority types. For instance, the < 2% of respondents who indicated that they do not have a choice about their Facebook use or non-use may be highly analytically interesting, for reasons described above. However, their severe sparsity in the data set makes them invisible in the typologies generated above. Such concerns have increasingly important implications when methods such as these are applied to the a broader range of typology generation tasks.

### 6.5 Implications for Quantitative Methodological Development

Our study has a number of implications for methodological development in HCI/CSCW/social computing. First, we adapted and implemented two different methods (curve-fitting and taxonomic-trees) for quantitative development of typologies that have been used with some success in other domains such as environmental engineering [44] and energy forecasting [61]. Our results broadly suggest that these methods can also be used in the generation of typologies around social data to operationalize complex, theoretical constructs. Prior work [22, 37, 40, 101] suggests that the overwhelming majority of work in typology development has been done through theory development and qualitative methods which don’t tend to work very well in building statistical models. Specifically, we allude to GLMs (i.e. linear regression, logistic regressions, multi-level models etc.) that are often built to classify typologies with specific sets of predictors. As the number of predicted classes (types) [72] increase, it becomes harder for the maximum likelihood estimator (MLE) algorithm upon which most GLMS are built to calculate an optimum solution. Thus, optimizing the number of predicted classes is an important consideration that purely qualitative methods may not be able to solve. Our approach aims to provide a balance between these two tradeoffs. In making this claim, we want to highlight the importance of qualitative methods in typology and theory development but also want to acknowledge that these results are often at odds with the current state of statistical practice [16, 26, 62, 81] in HCI/CSCW/social computing.

Second, our work can be placed in context with the existing state of statistical practices in HCI/CSCW/social computing. Broadly, quantitative researchers in the field either come from a social science background (with requisite expertise in inferential statistics and generalized linear models) or from a computer science background (with general expertise in machine learning) [26, 62].
Unfortunately, neither academic background lends itself well to rigorous statistical inference, which has rightfully caused researchers to raise concerns in how quantitative methods are developed and adapted in the community [16, 26, 62]. For instance, it is relatively well established that dependence on noisy p-values and Null Hypothesis Significance Testing (NHST) leads to research that is less statistically rigorous and replicable across different populations and contexts [62]. This is especially true when operationalizing complex theoretical constructs (such as in our case, non/use) for statistical modeling and inference. Our study suggests that there is space for deeper exploration of methodological development of statistically rigorous, yet accessible methods to the general HCI/CSCW/social computing researcher. For instance, one strong finding from our results suggests that complex theoretical constructs may be conceptualized through multiple predictive modeling approaches to yield greater confidence in patterns, trends and statistical inferences that actually exist in the data (i.e. true positives) and less confidence in those that do not (i.e. false positives or false negatives arising from randomness or noisiness of the data). This type of holistic modeling can be used to ascertain external validity of complex theoretical constructs across samples, populations and contexts beyond the current norm in the field.

Finally, in recent years, HCI/CSCW/social computing venues have been wrestling with the question of what constitutes methodological contribution to the field and whether there exists a set of methods unique to, or particularly amenable for, common questions that arise in HCI/CSCW/social computing [12, 62, 78, 81]. Our work contributes to the conversation in this space by expanding the domain of methodological contribution to typology development and by questioning the current state of statistical reasoning in examining complex, theoretical constructs. With the ever-changing nature of social media and non/use practices, statistical methods such as those presented in this paper may provide consistent, adaptable means to categorize and study these diverse phenomena.

7 ACKNOWLEDGMENTS

This work is supported in part by the US National Science Foundation (Grant No. IIS-1421498 and Grant No. CNS-1814533) and the Facebook Computational Social Science Methodology Research Award. Thanks to our survey participants for sharing their data and their experiences.

REFERENCES

[1] Alessandro Acquisti and Ralph Gross. 2006. Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In Proceedings of the Privacy Enhancing Technology Symposium. Springer, Cambridge, England, 36–58.

[2] Cecilie Schou Andreassen, Torbjørn Torsheim, Geir Scott Brunborg, and Ståle Pallesen. 2012. Development of a Facebook Addiction Scale. Psychological Reports 110, 2 (April 2012), 501–517. https://doi.org/10.2466/02.09.18.PR0.110.2.501-517

[3] Anne Archambault and Jonathan Grudin. 2012. A Longitudinal Study of Facebook, LinkedIn, & Twitter Use. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). Austin, TX, 2741–2750. https://doi.org/10.1145/2207676.2208671

[4] Eric Baumer, Mark Sueyoshi, and Bill Tomlinson. 2008. Exploring the Role of the Reader in the Activity of Blogging. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Florence, Italy, 1111–1120. https://doi.org/10.1145/1357054.1357228

[5] Eric P. S. Baumer. 2018. Socioeconomic Inequalities in the Non/Use of Facebook. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Montréal, QC, 616:1–616:14. https://doi.org/10.1145/3173574.3174190

[6] Eric P. S. Baumer, Phil Adams, Vera D. Khovanskaya, Tony C. Liao, Madeline E. Smith, Victoria Schwanda Sosik, and Kaiton Williams. 2013. Limiting, Leaving, and (Re)Lapsing: An Exploration of Facebook Non-Use Practices and Experiences. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Paris, 3257–3266. https://doi.org/10.1145/2470654.2466446

[7] Eric P. S. Baumer, Morgan G. Ames, Jenna Burrell, Jed R. Brubaker, and Paul Dourish. 2015. Why Study Technology Non-Use? First Monday 20, 11 (2015).
Methods for Generating Typologies of Non-use

[8] Eric P. S. Baumer and Jed R. Brubaker. 2017. Post-Userism. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Denver, CO, 6291–6303. https://doi.org/10.1145/3025453.3025740

[9] Eric P. S. Baumer, Jenna Burrell, Morgan G. Ames, Jed R. Brubaker, and Paul Dourish. 2015. On the Importance and Implications of Studying Technology Non-Use. interactions 22, 2 (2015), 52–56. https://doi.org/10.1145/2723667

[10] Eric P. S. Baumer, Shion Guha, Emily Quan, David Mimno, and Geri K. Gay. 2015. Missing Photos, Suffering Withdrawal, or Finding Freedom? How Experiences of Social Media Non-Use Influence the Likelihood of Reversion. Social Media + Society 1, 2 (2015). https://doi.org/10.1177/2056305115614851

[11] Eric P. S. Baumer, Shion Guha, Patrick Skeba, and Geraldine Gay. 2019. All Users are (Not) Created Equal: Predictors Vary for Different Forms of Facebook Non-use. Proceedings of the ACM on Human-Computer Interaction 3, CSCW (2019), 1–28.

[12] Eric P. S. Baumer, David Mimno, Shion Guha, Emily Quan, and Geri K Gay. 2017. Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? Journal of the Association for Information Science and Technology 68, 6 (2017), 1397–1410.

[13] Eric P. S. Baumer, Rui Sun, and Peter Schaedler. 2018. Departing and Returning: Sense of Agency As an Organizing Concept for Understanding Social Media Non/Use Transitions. Proc. ACM Hum.-Comput. Interact. 2, CSCW (Nov. 2018), 23:1–23:19. https://doi.org/10.1145/3274292

[14] Eric P. S. Baumer, Xiaotong Xu, Christine Chu, Shion Guha, and Geri K. Gay. 2017. When Subjects Interpret the Data: Social Media Non-Use as a Case for Adapting the Delphi Method to CSCW. In Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW) (CSCW ’17). ACM, Portland, OR, 1527–1543. https://doi.org/10.1145/2998181.2998182

[15] V. Benet-Martinez and O. P. John. 1998. Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. Journal of Personality and Social Psychology 75 (1998), 729–750.

[16] Michael S Bernstein, Mark S Ackerman, Ed H Chi, and Robert C Miller. 2011. The trouble with social computing systems research. In CHI’11 Extended Abstracts on Human Factors in Computing Systems. ACM, 389–398.

[17] Geoffrey C. Bowker and Susan Leigh Star. 1999. Sorting Things Out: Classification and Its Consequences. MIT Press.

[18] George EP Box and R Daniel Meyer. 1986. An analysis for unreplicated fractional factorials. Technometrics 28, 1 (1986), 11–18.

[19] danah boyd. 2006. Friends, Friendsters, and Top 8: Writing Community into Being on Social Network Sites. First Monday 11, 12 (Dec. 2006). https://doi.org/10.5210/fm.v11i12.1418

[20] danah boyd. 2014. It’s Complicated: The Social Lives of Networked Teens (1 edition ed.). Yale University Press, New Haven.

[21] Petter Bae Brandtzæg. 2010. Towards a unified Media-User Typology (MUT): A meta-analysis and review of the research literature on media-user typologies. Computers in Human Behavior 26, 5 (2010), 940–956.

[22] Petter Bae Brandtzæg and Jan Heim. 2011. A typology of social networking sites users. International Journal of Web Based Communities 7, 1 (2011), 28–51.

[23] Jed R. Brubaker, Mike Ananny, and Kate Crawford. 2016. Departing Glances: A Sociotechnical Account of ‘leaving’ Grindr. New Media & Society 18, 3 (2016), 373–390. https://doi.org/10.1177/146144481542311

[24] Jed R. Brubaker and Vanessa Callison-Burch. 2016. Legacy Contact: Designing and Implementing Post-Mortem Stewardship at Facebook. In Proc CHI. ACM Press, 2908–2919. https://doi.org/10.1145/2858036.2858254

[25] Tom Buchanan, Carina Paine, Adam N. Joinson, and Ulf-Dietrich Reips. 2007. Development of Measures of Online Privacy Concern and Protection for Use on the Internet. Journal of the American Society for Information Science and Technology 58, 2 (2007), 157–165. https://doi.org/10.1002/asi.20459

[26] Paul Cairns. 2007. HCI... not as it should be: inferential statistics in HCI research. In Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 1. British Computer Society, 26–29.

[27] Petter Bae Brandtzæg and Jan Heim. 2011. A typology of social networking sites users. International Journal of Web Based Communities 7, 1 (2011), 28–51.

[28] Yushi Chen, Zhouhan Lin, Xing Zhao, Gang Wang, and Yanfeng Gu. 2014. Deep learning-based classification of hyperspectral data. IEEE Journal of Selected topics in applied earth observations and remote sensing 7, 6 (2014), 2094–2107.

[29] Paul T. Costa and Robert R. McCrae. 1992. Revised NEO Personality Inventory (NEO-PIR) and NEO Five Factor Inventory (NEO-FFI) Professional Manual. Psychological Assessment Resources, Odessa, FL.

[30] Jenny Davis. 2012. The Problem with Internet Addiction. Psychological Assessment Resources, Odessa, FL.

[31] Thomas G Dieterich. 2000. Ensemble methods in machine learning. In International workshop on multiple classifier systems. Springer, 1–15.

[32] Eric P. S. Baumer, Rui Sun, and Peter Schaedler. 2018. Departing and Returning: Sense of Agency As an Organizing Concept for Understanding Social Media Non-Use Transitions. Proc. ACM Hum.-Comput. Interact., Vol. 4, No. CSCW1, Article 27. Publication date: May 2020.
[33] Maeve Duggan, Nicole B. Ellison, Cliff Lampe, Amanda Lenhart, and Mary Madden. 2015. *Pew Social Meida Report 2015*. Technical Report January. Pew Research Center, Washington, D.C.

[34] Susan Dumais, Robin Jeffries, Daniel M. Russell, Diane Tang, and Jaime Teven. 2014. Understanding User Behavior Through Log Data and Analysis. In *Ways of Knowing in HCI* Judith S. Olson and Wendy A. Kellogg (Eds.), Springer New York, New York, NY, 349–372. https://doi.org/10.1007/978-1-4939-0378-8_14

[35] Allen Louis Edwards. 1983. *Techniques of Attitude Scale Construction*. Prentice-Hall, Englewood Cliffs, NJ.

[36] Nicole B. Ellison, Charles Steinfield, and Cliff Lampe. 2007. The Benefits of Facebook "Friends:" Social Capital and College Students’ Use of Online Social Network Sites. *Journal of Computer-Mediated Communication* 12, 4 (July 2007), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

[37] Chris Eldsen, Arti Manohar, Jo Briggs, Mike Harding, Chris Speed, and John Vines. 2018. Making sense of blockchain applications: A typology for HCI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 458.

[38] Facebook. [n. d.]. How Do I Deactivate My Account? https://www.facebook.com/help/214376678584711.

[39] Gloria Fraser. 2018. Evaluating inclusive gender identity measures for use in quantitative psychological research. *Psychology & Sexuality* 9, 4 (2018), 343–357.

[40] Christian Fuchs and Marianna Obrist. 2010. HCI and society: towards a typology of universal design principles. *Int'l Journal of Human–Computer Interaction* 26, 6 (2010), 638–656.

[41] Mike Furr. 2011. *Scale Construction and Psychometrics for Social and Personality Psychology*. SAGE Publications Ltd, London.

[42] Radhika Garg. 2019. An analysis of (non-) use practices and decisions of internet of things. In *IFIP Conference on Human-Computer Interaction*. Springer, 3–24.

[43] Radhika Garg and Subhasree Sengupta. 2019. "When you can do it, why can’t I?": Racial and Socioeconomic Differences in Family Technology Use and Non-Use. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–22.

[44] Lance Gentry, Roger J Calantone, and Shaojie Anna Cui. 2006. The forecasting classification grid: a typology for method selection. *Journal of Global Business Management* 2, 1 (2006), 48–60.

[45] Ilana Gershon. 2011. Un-Friend My Heart: Facebook, Promiscuity, and Heartbreak in a Neoliberal Age. *Anthropological Quarterly* 84, 4 (Nov. 2011), 865–894. https://doi.org/10.1353/anq.2011.0048

[46] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. 2003. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality* 37, 6 (Dec. 2003), 504–528. https://doi.org/10.1016/S0092-6566(03)0046-1

[47] Sukeshini A Grandhi, Linda Plotnick, and Starr Roxanne Hiltz. 2019. Do I Stay or Do I Go? Motivations and Decision Making in Social Media Non-use and Reversion. *Proceedings of the ACM on Human-Computer Interaction* 3, GROUP (2019), 1–27.

[48] Jon E. Grant. 2008. *Impulse Control Disorders: A Clinician’s Guide to Understanding and Treating Behavioral Addictions*. W. W. Norton & Company, Inc., New York.

[49] Jon E. Grant, Marc N. Potenza, Aviv Weinstein, and David A. Gorelick. 2010. Introduction to Behavioral Addictions. *American Journal of Drug and Alcohol Abuse* 36, 5 (Aug. 2010), 233–241. https://doi.org/10.3109/00952990.2010.491884

[50] Mark Griffiths. 2005. A ìAAcomponentsî model of addiction within a biopsychosocial framework. *Journal of Substance Use* 10, 4 (Jan 2005), 191–197. https://doi.org/10.1080/14659900500114359

[51] Mark D. Griffiths. 2013. Social Networking Addiction: Emerging Themes and Issues. *Journal of Addiction Research & Therapy* 4, 5 (2013). https://doi.org/10.4172/2155-6105.1000e118

[52] Mark D. Griffiths, Daria J. Kuss, and Zsolt Demetrovics. 2014. Social Networking Addiction. In *Behavioral Addictions*. Elsevier, 119–141.

[53] Shion Guha, Eric P. S. Baumer, and Geri K. Gay. 2018. Regrets, I’ve Had a Few: When Regretful Experiences Do (and Don’t) Compel Users to Leave Facebook. In *Proceedings of the ACM Conference on Supporting Group Work (GROUP)*. ACM, Sanibel Island, FL, 166–177. https://doi.org/10.1145/3148330.3148338

[54] Eszter Hargittai. 2008. Whose Space? Differences Among Users and Non-Users of Social Network Sites. *Psychology & Sexuality* 2, 1 (2006), 48–60.

[55] Eszter Hargittai and Yu-li Patrick Hsieh. 2010. From Dabblers to Omnivores: A Typology of Social Network Site Usage. *Journal of Computer-Mediated Communication* 15, 1 (2008), 276–297. https://doi.org/10.1111/j.1083-6101.2007.00396.x

[56] Ellie Harmon and Melissa Mazmanian. 2013. Stories of the Smartphone in Everyday Discourse: Conflict, Tension & Instability. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. Paris, 1051–1060. https://doi.org/10.1145/2470654.2466134

[57] Samantha Jaroszewski, Danielle Lotttridge, Oliver L Haimson, and Katie Quehl. 2018. "Genderfluid” or “Attack Helicopter” Responsible HCI Research Practice with Non-binary Gender Variation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–15.
Methods for Generating Typologies of Non/use 27:23

[58] Yongick Jeong and Yeuseung Kim. 2017. Privacy concerns on social networking sites: Interplay among posting types, content, and audiences. Computers in Human Behavior 69 (2017), 302–310. https://doi.org/10.1016/j.chb.2016.12.042

[59] O. P. John, E. M. Donahue, and R. L. Kentle. 1991. The Big Five Inventory - Versions 4a and 54. Technical Report. University of California Berkeley, Institute of Personality and Social Research, Berkeley, CA.

[60] O. P. John, L. P. Naumann, and C. J. Soto. 2008. Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In Handbook of personality: Theory and research, O. P. John, R. W. Robins, and L. A. Pervin (Eds.). Guilford Press, New York, NY, 114–158.

[61] Dae-Ki Kang and Kiwook Sohn. 2009. Learning decision trees with taxonomy of propositionalized attributes. Pattern Recognition 42, 1 (2009), 84–92.

[62] Matthew Ka V, Steve Haroz, Shion Guha, and Pierre Dragicevic. 2016. Special interest group on transparent statistics in hci. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 1081–1084.

[63] Funda Kivran-Swaine, Priya Govindan, and Mor Naaman. 2011. The Impact of Network Structure on Breaking Ties in Online Social Networks: Unfollowing on Twitter. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Vancouver, BC, 1101–1104.

[64] Haewoon Kwak, Hyunwoo Chun, and Sue Moon. 2011. Fragile Online Relationship: A First Look at Unfollow Dynamics in Twitter. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Vancouver, BC, 1091–1100. https://doi.org/10.1145/1978942.1979104

[65] Cliff Lampe, Jessica Vitak, and Nicole Ellison. 2013. Users and Nonusers: Interactions between Levels of Facebook Adoption and Social Capital. In Proceedings of the ACM Conference on Computer Supported Cooperative Work ( CSCW). Baltimore, MD, 809–819. https://doi.org/10.1145/2441776.2441867

[66] Alex Leavitt. 2015. “This Is a Throwaway Account”: Temporary Technical Identities and Perceptions of Anonymity in a Massive Online Community. In Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing ( CSCW ’15). ACM, Vancouver, BC, 317–327. https://doi.org/10.1145/2675133.2675175

[67] Dominik J. Leiner, Lara Kobilke, Christina Rueß, and Hans-Bernd Brosius. 2018. Functional Domains of Social Media Platforms: Structuring the Uses of Facebook to Better Understand Its Gratifications. Computers in Human Behavior 83 (June 2018), 194–203. https://doi.org/10.1016/j.chb.2018.01.042

[68] Amanda Lenhart and John B. Horrigan. 2003. Re-Visualizing the Digital Divide as a Digital Spectrum. IT & Society 1, 5 (2003), 23–39.

[69] Hanlin Li, Nicholas Vincent, Janice Tsai, Jofish Kaye, and Brent Hecht. 2019. How Do People Change Their Technology Use in Protest? Understanding. Proceedings of the ACM on Human-Computer Interaction 3, CSCW (2019), 1–22.

[70] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. 2018. What Makes Smartphone Use Meaningful or Meaningless? Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 1 (March 2018), 22:1–22:26. https://doi.org/10.1145/3191754

[71] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. 2019. Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI). ACM, Glasgow, 131:1–131:18. https://doi.org/10.1145/3290605.3300361

[72] Henrik Madsen and Pouil Thyregod. 2010. Introduction to general and generalized linear models. CRC Press.

[73] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. 2004. Internet Users’ Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. Information Systems Research 15, 4 (2004), 336–355.

[74] Alice E. Markwick and danah boyd. 2014. Networked Privacy: How Teenagers Negotiate Context in Social Media. New Media & Society 16, 7 (Nov. 2014), 1051–1067. https://doi.org/10.1177/1461444814536513

[75] Robert R. McCrae and Paul T. Costa. 1997. Personality Trait Structure as a Human Universal. American Psychologist 52, 1 (Jan. 1987), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

[76] Robert R. McCrae and Paul T. Jr. Costa. 1997. Personality Trait Structure as a Human Universal. American Psychologist 52, 5 (May 1997), 509–516. https://doi.org/10.1037/0003-066X.52.5.509

[77] Michael Muller, Shion Guha, Eric P. S. Baumer, David Mimno, and N Sadat Shami. 2016. Machine learning and grounded theory method: Convergence, divergence, and combination. In Proceedings of the 19th International Conference on Supporting Group Work. ACM, 3–8.

[78] Blair Nonnecke and Jenny Preece. 2001. Why Lurkers Lurk. In Proceedings of the Americas Conference on Information Systems. Boston.

[79] Daniel Oberski. 2016. Mixture models: Latent profile and latent class analysis. In Modern statistical methods for HCI. Springer, 275–287.

[80] Judith S Olson and Wendy A Kellogg. 2014. Ways of Knowing in HCI. Vol. 2. Springer.
A  CORRELATION MATRIX FOR NON/USE QUESTIONS

| Question                          | Typology-1 | Typology-2 | Typology-3 | Typology-4 | Typology-5 | Typology-6 | Typology-7 |
|-----------------------------------|------------|------------|------------|------------|------------|------------|------------|
| Typology-1                        | 1.00       | 0.09       | -0.19      | -0.33      | -0.04      | 0.00       | -0.17      |
| Typology-2                        | 0.09       | 1.00       | 0.12       | 0.06       | 0.04       | 0.02       | 0.07       |
| Typology-3                        | -0.19      | 0.12       | 1.00       | 0.45       | 0.37       | 0.16       | 0.55       |
| Typology-4                        | -0.33      | 0.06       | 0.45       | 1.00       | 0.22       | 0.24       | 0.36       |
| Typology-5                        | -0.04      | 0.04       | 0.37       | 0.22       | 1.00       | 0.20       | 0.42       |
| Typology-6                        | 0.00       | 0.02       | 0.16       | 0.24       | 0.20       | 1.00       | 0.23       |
| Typology-7                        | -0.17      | 0.07       | 0.55       | 0.36       | 0.42       | 0.23       | 1.00       |

**Typology-1:** I currently have a Facebook account  
**Typology-2:** I have more than one Facebook account  
**Typology-3:** At some point, I have deactivated my Facebook account  
**Typology-4:** At some point, I have permanently deleted my Facebook account  
**Typology-5:** At some point, I have voluntarily taken a break from Facebook for a week or more  
**Typology-6:** At some time, I have used software to limit my Facebook usage  
**Typology-7:** At some point, I have deleted the Facebook app from my phone  

Table 5. Correlation matrix for non/use questions

B  FACTOR LOADINGS

The following tables show all questions and factor loadings used to derive the psychometric predictors in Section 4.2.

| Question                                      | Salience, Tolerance & Mood (α = .83) | Withdrawal & Conflict (α = .84) | Relapse |
|-----------------------------------------------|--------------------------------------|---------------------------------|---------|
| Spent a lot of time thinking about FB use     | 0.83                                 |                                 |         |
| Felt an urge to use FB more and more          | 0.85                                 |                                 |         |
| FB use to forget about personal problems      | 0.57                                 |                                 | 0.35    |
| Became restless if FB use prohibited          |                                      | 0.96                            |         |
| FB use had a negative impact on job/studies   |                                      | 0.97                            |         |
| Tried to cut down FB use without success      |                                      |                                 | 0.84    |

Table 6. BFAS: Factor analysis.

| Question                                      | Connectedness (α = .86) | Daily Routine (α = .93) | Pride |
|-----------------------------------------------|-------------------------|-------------------------|-------|
| I feel out of touch when I haven’t logged onto FB for a while | 0.64                    | 0.35                    |       |
| I feel I am part of the FB community          | 0.66                    | 0.39                    | 0.38  |
| I would be sorry if FB shut down              | 0.71                    | 0.35                    | 0.30  |
| FB is part of my everyday activity            | 0.41                    | 0.86                    |       |
| FB has become part of my daily routine        | 0.58                    | 0.65                    |       |
| I am proud to tell people I am or was on FB   | 0.35                    |                         | 0.89  |

Table 7. FBI: Factor analysis
| Question                                                                 | Sense of Negative Agency | Sense of Positive Agency |
|-------------------------------------------------------------------------|--------------------------|--------------------------|
|                                                                         | ($\alpha = .84$)         | ($\alpha = .79$)         |
| I am just an instrument in the hands of something else                  | 0.57                     | -0.21                    |
| My actions just happen without my intention                             | 0.70                     |                          |
| Consequences of my actions don’t logically follow my actions            | 0.71                     |                          |
| My movements are automatic - my body simply makes them                   | 0.64                     |                          |
| The outcomes of my actions generally surprise me                         | 0.75                     |                          |
| Nothing I do is actually voluntary                                      | 0.50                     |                          |
| I feel like I am a remote-controlled robot                               | 0.68                     |                          |
| I am in full control of what I do                                       |                          | 0.60                     |
| I am the author of my actions                                           | -0.23                    | 0.74                     |
| Things I do are subject only to my free will                            |                          | 0.55                     |
| The decision whether and when to act is within my hands                  |                          | 0.63                     |
| My behavior is planned by me from the beginning to the end              |                          | 0.55                     |
| I am responsible for everything that results from my actions             |                          | 0.67                     |

Table 8. Sense of Agency: Factor analysis

| Question                                                                 | Personal integration ($\alpha = .84$) | Social integration ($\alpha = .91$) | Escape ($\alpha = .81$) | Affective Gratification ($\alpha = .91$) |
|-------------------------------------------------------------------------|----------------------------------------|-------------------------------------|-------------------------|------------------------------------------|
| I use FB because it makes me ease off                                   | 0.50                                   | 0.31                                |                          |                                          |
| I use FB to inform myself about certain topics                          | 0.57                                   | 0.31                                | 0.31                    |                                          |
| I use FB to receive advice and recommendations                          | 0.61                                   |                                     |                          |                                          |
| I use FB to express who I am                                           | 0.72                                   |                                     |                          |                                          |
| I use FB to share my views and opinions                                 | 0.73                                   |                                     |                          |                                          |
| I use FB to keep in touch with friends                                  | 0.89                                   |                                     |                          |                                          |
| I use FB to exchange with my friends and family                         | 0.31                                   | 0.70                                |                          | 0.68                                     |
| I use FB because I am bored                                             | 0.31                                   |                                     | 0.68                    |                                          |
| I use FB to occupy myself                                               | 0.39                                   | 0.35                                | 0.61                    |                                          |
| I use FB because it is fun                                               | 0.34                                   | 0.33                                | 0.84                    |                                          |

Table 9. Leiner’s Scale: Factor analysis