An optimization framework to provide volunteers with task selection autonomy and group opportunities

Milan Preet Kaur[1], Jennifer A. Pazour[1]*, Rosemonde Ausseil[2]

[1]*Rensselaer Polytechnic Institute, Industrial and Systems Engineering Department, 110 8th Street, Troy, NY, 12180, USA*

[2]*Clarkson University, Reh School of* Business, *34 Wahta Dr, Potsdam, NY, 13699, USA*

*Corresponding Author, pazouj@rpi.edu

Abstract

Nonprofit Organizations (NPOs) rely on volunteers to support community needs but struggle with making strategic volunteer-to-task assignments to enable volunteer satisfaction, and completion of complex tasks. Creation of volunteer groups and their assignment to NPO tasks can help achieve these goals by providing volunteers with opportunity for networking, collaboration, and peer learning. However, strategically creating ideal assignments is challenging because (i) there are exponentially many ways a set of volunteers can be assigned in groups; and (ii) NPOs tend to have limited and uncertain data concerning volunteers' personal preferences, availabilities, and motivations to participate. To address these challenges, this research contributes by introducing an integer programming framework to offer volunteers a menu of tasks to choose from and then based on volunteers' willingness information, creates ideal homogenous volunteer group assignments. These groups are created such that the group collectively meet a task's skill requirements and groups of volunteers of similar skill and affinity levels are prioritized. We apply the developed methodology to a case study based on a partner NPO that works with remote volunteers from multiple countries to produce online educational content. The menu creation method can improve NPO and volunteer-based performance metrics, where the most improvement is observed when a NPO is faced with very picky volunteers. Presenting volunteers with larger menus of tasks also leads to an improvement in ideal group creations. Implementing the group creation methodology helps obtain a statistically significant increase in ideal group creations but results in a tradeoff of decreased benefits to volunteers and the NPO. Finally, implementing a minimum desired group size does not severely impact most KPIs and would be beneficial for an NPO to implement as it encourages the creation and assignment of volunteer groups to tasks.

Keywords: volunteer crowdsourcing platform, static integer programs, task recommendation menu creation, homogenous group creation, learning, affinity, volunteer satisfaction, volunteer motivations, volunteer group to task assignments

1. Introduction

Humanitarian Nonprofit Organizations (NPOs), which represent 35.2% of all US NPOs (NCCS Report, 2022), provide essential community services, such as housing, education, healthcare, food, and disaster relief. To provide these services, NPOs rely heavily on volunteers. Thus, a critical volunteer management decision for NPOs is how to make volunteer-to-task assignments that can balance the NPO's project goals with their volunteers' motivations to contribute (*Deloitte Volunteer Impact Research*, 2017; Lo et al., 2021). This is challenging because volunteers are not employed by an NPO, and so an NPO cannot dictate to volunteers which task they must complete. As these volunteers have varying personal motivations to volunteer and possess varying skill levels, making strategic volunteer-to-task assignments influences both task performance and volunteer satisfaction (Berenguer & Shen, 2020; Berenguer et al., 2023; Millette & Gagné, 2008). Qualitative literature highlights that people volunteer due to personal intrinsic factors, but task and environmental factors, such as autonomy in task selection, skill learning opportunities, personal development, knowledge application, networking, and building interpersonal relationships, also significantly impact volunteer satisfaction (Alam & Campbell, 2017; Barnes & Sharpe, 2009; Barron & Rihova, 2011; Baruch et al., 2016; Fullwood & Rowley, 2021; Mazlan et al., 2018; Millette & Gagné, 2008).

Moreover, observations of volunteers working in groups highlight the positive impact of knowledge sharing and affinity levels among volunteers on organizational needs and individual engagement and performance (Fait & Sakka, 2020; Fullwood & Rowley, 2021; Jensen & McKeage, 2015). Specifically, collaborative learning, where peers exchange knowledge, information, and skills relevant to the task at hand, can be accomplished in in-person and online tasks by assigning volunteer groups to work together. A group setting can lead to a deeper level of learning, critical thinking, and voluntary long-term engagement of individuals towards completing their work while providing benefits like structured and meaningful social interaction, which aids the development of social relationships among participating individuals (Zheng et al., 2018). In addition, assigning groups of volunteers to a task also benefits the NPO as more complex tasks can be tackled by volunteers through their collective skills and abilities. However, making strategic volunteer group-to-task assignments is challenging because (i) there are exponentially many ways a set of volunteers can be assigned in groups; and (ii) NPOs tend to have limited access to volunteer data or low-quality data concerning volunteers' personal preferences, availability, and skills (Fullwood & Rowley, 2021; Jensen & McKeage, 2015).

To address these challenges, we design a methodology that creates volunteer groups and assigns them to tasks. A key aspect of our methodology is it first presents volunteers with a menu of possible tasks, from which the volunteers are asked to select which of tasks they are willing to contribute to. This menu-based group formation approach is beneficial as it enables a NPO to:

1. Create and offer a menu of personalized task recommendations to provide volunteers with autonomy in selecting tasks they wish to work on.
2. Take a systematic view of tasks and volunteers to make strategical volunteer group to task assignments that balance the task completion needs of the NPO along with volunteers' preferences, skills, and motivations.
3. Create volunteer group to task assignments even when NPOs have limited volunteer data. This approach is useful when an NPO has volunteer information (e.g. demographic information) that can be used to estimate volunteer's task preferences. These estimates are used to create personalized menu of tasks, but also acknowledges that these estimates are imperfect and hedges against this estimation uncertainty by allowing volunteers to indicate their willingness to

volunteer on multiple task options, and a volunteer is assigned to a task only if they have stated their willingness to work on it.

4. Create ideal volunteer groups and their assignment to tasks, which can aid a NPO in successful completion of complex tasks. It also supports the varying personal motivations of volunteers that can be achieved through group-work (Alam & Campbell, 2017; Barnes & Sharpe, 2009; Barron & Rihova, 2011; Fait & Sakka, 2020; Fullwood & Rowley, 2021; Jensen & McKeage, 2015; Wiens et al., 2022).

This research contributes by creating a new optimization framework that provides volunteers with task recommendation menus to select from and then strategically forms volunteer groups and assigns them to tasks. The menu-creation method determines volunteer menus, considering the need to form ideal groups and captures uncertainty in volunteers' willingness behavior by combining solutions from solving a set of deterministic integer program for different volunteer willingness scenarios. Volunteer task selections from their presented menus are then collected. Then, the group creation model, an integer programming model, creates volunteer-to-task assignments that maximizes the total number of volunteer to task assignments made and considers creating ideal volunteer groups (with best possible learning and affinity levels within groups) and assigns them to appropriate tasks. We partnered with an NPO that produces online educational content and conducted an empirical data analysis from their initial onboarding form offered to volunteers to describe the probability distribution of volunteers in different age, conditional probability distributions of volunteer's task selections based on their age, probability distribution of number of task selections volunteers made on the onboarding form, and the probability distribution of volunteers with different skill levels in each skill defined for the completion of online tasks. These distributions are used to conduct a computational experiment, which indicates that our proposed methodology can benefit both NPOs and volunteers alike.

The rest of the paper is organized as follows. In Section 2, we review related literature. We then describe our methodology and model formulation in Section 3. Sections 4 and 5 contain a detailed description of the empirical data analysis of the data obtained from the partner NPO's database, and a detailed description of our design of experiments based on a volunteer crowdsourcing platform completing complex online tasks. In Sections 6 and 7, we present our results, conclusions and future research ideas.

2. Literature Review

We review literature related to: (1) volunteer to task assignment methodologies, (2) menu creation methods for worker to task assignment in for-profit applications, and (3) group creation and task assignment approaches.

2.1 Volunteer to Task Assignment Methodologies

Researchers have designed various volunteer-to-task assignment methodologies and these methodologies can be classified into three types: The **Server Assigned Tasking (SAT)** is a centralized method that uses volunteer and task information available to the organization to make volunteer-to-task assignments and is currently the most studying volunteer management method, but does not provide volunteers with autonomy to choose tasks they are willing to help with (Kazemi, 2012). Most of these papers use deterministic integer programming approaches for resource allocation to tasks while accounting for (assumed known) volunteer preferences and availabilities (Garcia et al., 2018; Kaur et al., 2022; Rauchecker & Schryen, 2018). The **Worker Selected Tasking (WST)** method is a decentralized volunteer-to-task assignment approach where volunteers select a preferred task from a list of all active tasks (Kazemi, 2012). This approach (e.g., (Henderson et al., 2022) prioritizes individual volunteer preferences by allowing volunteers full autonomy in task selection but causes search friction due to the presentation of all active tasks to the volunteer. This method does not incorporate organizational preferences and therefore can lead to task assignments that are systematically inefficient for the NPO. The WST method is commonly employed on volunteer crowdsourcing platforms that match volunteers with either in-person or online tasks. The **Menu Creation** (MC) method is a hybrid of the SAT and WST methods and consists of a three-step process: First, the NPO incorporates relevant individual and organization preferences to present a smaller list (menu) of active tasks specific to each active volunteer. Next, a volunteer selects their most preferred task(s) from the task menu offered to them. Lastly, these volunteer task selections are collected, and assignments are made that meet applicable constraints (e.g., so that maximum number of volunteers are assigned to tasks). The MC method allows for strategic considerations of the NPO's goals while still incorporating volunteer's many preferences in the task assignment process. Further, this approach is useful when a NPO has an estimate of volunteer information, but not a perfect estimate of what a volunteer's preferences would look like when they arrive at the organization to volunteer.

Volunteer management literature deploying a Menu Creation approach is limited. Schmidt et al. introduces a dynamic integer program using the Menu Creation method and generates a menu of ranked task recommendations for each volunteer (Schmidt & Albert, 2022). Task ranking is based on features identified to measure volunteer and organizational preferences and their objective function maximizes the posting of the most preferable tasks on each volunteer's menu of tasks. Volunteers then select a task from the menu of tasks presented to them. Their model assumes that the higher ranking (position) of tasks on a menu recommendation to a volunteer increases the chances that a task gets picked. Their model is designed to implement one-to-one volunteer-to-task assignments and is designed to support disaster relief services. Escallon-Barrios et al. introduce a dynamic integer program to offer volunteers a menu of multiple time slots on online volunteer crowdsourcing platforms (Escallon-Barrios et al., 2024). A central planner first assigns employees to time slots, which limits the time slots that can be presented to volunteers (only one of either employee or volunteer is assigned per time slot), in an attempt to influence volunteer slot selections (Escallon-Barrios et al., 2024).

Literature developing a related methodology called Notification Mechanisms are also reviewed, where models are designed with the aim to regulate notifications sent to volunteers (Henderson et al., 2022; Lo et al., 2021; Manshadi & Rodilitz, 2021; Shi et al., 2021a; Tongarlak et al., 2024). These papers create methodologies for notification policy designs, which focus on notification decisions rather than assignment decisions (i.e., these notification papers do not capture individual volunteer to task assignments nor volunteer task selections). Chen et al., 2021 & Chen et al., 2021 design models to support self-selected volunteer group dispatching by developing a Menu Creation methodology for volunteer management (Chen, Wang, et al., 2021; Chen, Zhang, et al., 2021). They develop a two-sided static integer programming methodology that incorporates the information of two agents: self-selected volunteer groups and the service demanders (community members) where the service demanders have varying needs and preferences for which group they select to work on their task. They utilize the initially collected preferences to recommend a menu of options to both agents. Each agent then selects item(s) on their menus and provides the platform with their revised preferences. Their model aims to achieve maximum satisfaction for volunteer groups and task requestors and is designed to reduce the uncertainty in creating matches between the two agents with clashing preferences. Their models are designed to support a crowdsourcing platform that enables multiple organizations to post their task requirements and is open to all volunteers visiting the site. These models do not cater to individual volunteer arrivals and creation of groups and their assignment to tasks.

2.2 Menu Creation Methods for Worker to Task Assignment in For-Profit Applications

Given the limited literature developing the Menu Creation method for NPO applications, we also review optimization models designed according to the menu creation methodology that make worker-to-task assignments in for profit settings where workers also are independent of the organization they serve (e.g., crowdsourced delivery and ridesharing). A systematic literature review on worker crowdsourcing and task assignment methods on crowdsourcing platforms identify the Menu Creation method as a recently developed and promising methodology ideal for achieving various organizational (successful task completion, worker retention, minimal resource and funds usage, etc.) and worker goals (skill learning, networking, reduced task selection friction, increased autonomy in task selection etc.) (Zhen et al., 2021). The Menu Creation method has been developed for peer-to-peer logistics services (e.g., crowdsourced delivery and ride sharing) where workers are independent contractors. Mofidi et al., 2019 studies the creation of personalized task recommendations made to workers who then select their most preferred task. Assuming deterministic task selection behavior, they create a bi-level deterministic optimization model (Mofidi & Pazour, 2019). Horner et al., 2021 furthers this research to capture stochastic selection behaviors by developing a Sample Average Approximation technique where workers can inform their willingness to work on tasks recommended on a menu, which is then used to make final worker to task assignments (Horner et al., 2021). Ausseil et al., 2022, extends this work to a dynamic setting, in which both tasks and workers arrive spontaneously over time (Ausseil et al., 2022). Notably, these works implement the Menu Creation method to create task menus that narrows down options presented to the workers on the platform and make one-to-one worker (driver) to task assignments.

2.3 Group Creation and Task Assignment Approaches

Group Creation and Assignment approaches create groups from a set of independently arriving volunteers and assigns the groups to tasks based on the preferences of the volunteers as well as the NPO's task requirements that a created group can meet. Yet, in the nonprofit volunteer assignment literature only self-selected groups have been considered, meaning the groups arrive together and are assigned together (Chen, Wang, et al., 2021; Chen, Zhang, et al., 2021; Falasca & Zobel, 2012).

The implementation of different group creation approaches has been widely studied in educational and for-profit applications and we review literature that creates optimization models to create groups, focusing on models with various objectives such as interpersonal compatibility, social network connectivity, and learning potential. Rahman et al., 2019 develop a static integer programming model, where they collect and incorporate worker information such as skills, availability, wages, etc. to create worker groups ideal for increased collaboration through the consideration of affinity among group members while satisfying organization's task's skill and cost requirements (Rahman et al., 2019). Vinella et al., 2022 develop a dynamic algorithm for the group creation of crowdsourced workers hired for completion of complex tasks (Vinella et al., 2022). They implement similar approaches as Rahman et al., 2019 but also incorporate changing worker preferences throughout the project timeline through the dynamic nature of their model. Esfandiari et al., 2019 develop an integer programming methodology to aid group creation and implement heterogeneity in skill levels (maximizing skill level differences among group members) to achieve maximum learning and implement homogeneity in affinity levels (similarity in demographics and psychological traits) to support successful collaborations within groups (Esfandiari et al., 2019). Wei et al., 2021 develop a dynamic algorithm to enhance peer learning through targeted group creation by measuring aggregate learning gain per individual to obtain maximum learning, through the creation of heterogenous groups (Wei et al., 2021). Minn et al., 2018 developed a model to track student learning level and groups students dynamically based on progress made in previous periods into homogenous groups (Minn et al., 2018). They assess at each time interval, the student's improvement in knowledge along with their learning ability in particular skills to put them into groups to help predict a student's performance while tracing their learning information dynamically (Minn et al., 2018). Sanchez et al., 2021 create a model to create homogeneous groups in collaborative learning environments (education) and create groups based on student personality traits (affinity) for better academic performance (Sánchez et al., 2021). Most of the literature developing models creating heterogenous groups was aiming for groups ideal for a classroom setting where the maximization of difference in skill is considered the ideal approach to achieve fair student groups with equal learning opportunities (Esfandiari et al., 2019; Wei et al., 2021; Zheng et al., 2018). Some papers implemented the formation of homogenous groups and highlighted respective benefits (Minn et al., 2018; Rahman et al., 2019; Sánchez et al., 2021; Vinella et al., 2022; Zheng et al., 2018).

We determine homogenous group creations as ideal for NPO applications as our goal is to strengthen volunteer's personal motivations, provide them with the opportunity to engage in gradual and consistent learning of skills, and give them a sense of responsibility and independence in their volunteering involvement.

2.4 Summary

As illustrated in Figure 1, our work contributes to the domain of making volunteer to task assignments in NPO applications and integrates two approaches that to the best of our knowledge have not been combined previously: menu creation method with group creation and task assignment.

To our knowledge, only two recent works develop the Menu Creation methodology to aid volunteer to task assignments on volunteer crowdsourcing platforms (Escallon-Barrios et al., 2024; Schmidt & Albert, 2022) and both focus on one-to-one volunteer to task assignments. Thus, this work builds upon this literature by developing a menu creation methodology that:

1. Offers volunteers with task recommendation menus to select preferred tasks from
2. Creates volunteer groups (from individual arriving volunteers) and assigns them to NPO tasks
3. Integrates the menu creation and group creation methodologies

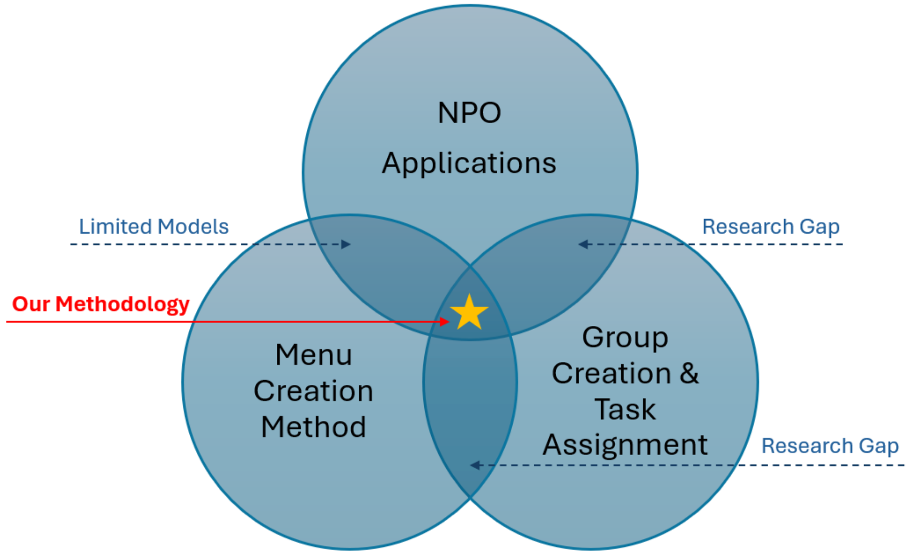


Figure 1: Contributions of this research paper and gaps in existing literature

3. Methodology Description

In this section we present a methodology to create and offer personalized task recommendation menus for volunteers to choose from, and then based on volunteers' responses on willingness to work on task(s), create and assign ideal volunteer groups to NPO tasks. The methodology incorporates the motivations and needs of both NPOs and volunteers and is designed for NPOs that store basic volunteer information collected during volunteer onboarding. As shown in Figure 2, this methodology assumes a NPO has a set of tasks, with respective skill requirements, and access to a set of volunteers the NPO can notify to inquire about their willingness and availability to help with these tasks.

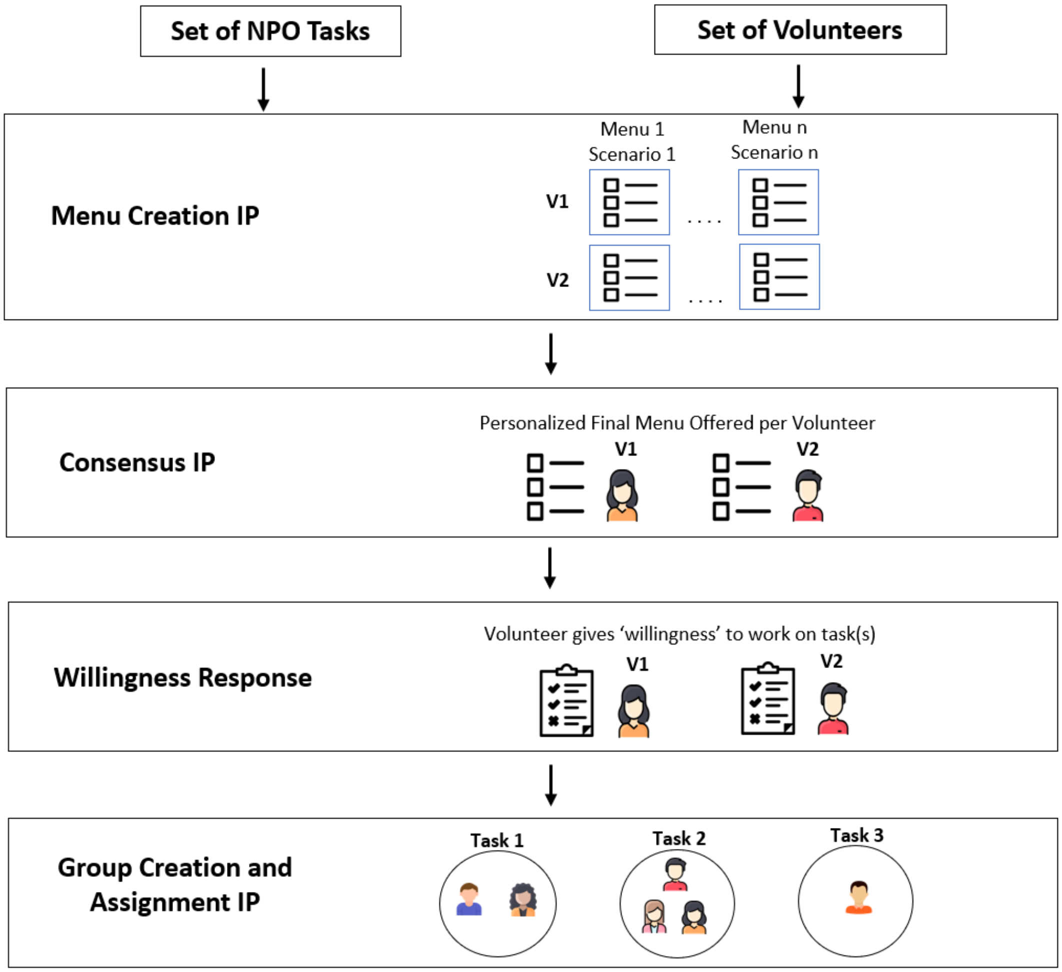


Figure 2: Process flow diagram for the Menu Creation and Group Creation and Assignment volunteer management method

Our methodology consists of solving three integrated integer programs, one to create personalized volunteer menus (MC Method), one to determine the consensus of the menus generated across multiple scenarios, and one for group creation and task assignments (GC Method).

First, the MC method is designed to create personalized task recommendation menus for each active volunteer. A menu is a subset of the active tasks in the system. The MC method is designed to holistically manage volunteer's task selections to avoid selection of highly preferred tasks by many volunteers, while ignoring the less popular tasks at an NPO. This is implemented by introducing a menu size that limits the number of tasks recommended to each volunteer. Narrowing down which tasks are presented to each volunteer helps tackle the issue of task selection friction commonly faced by volunteers. Tasks in different volunteer menus may be duplicated (i.e., a task may be offered to multiple volunteers simultaneously). This is an important feature offered by the MC method because having a task on multiple lists increases the chances of a task finding volunteers willing to help on them. This method also provides volunteers with autonomy in task selection while decreasing the NPO's need for accurate volunteer information to make volunteer management decisions as the methodology collects volunteer's task selections from menus being offered. Our method uses estimates of volunteer information available from NPO's volunteer on-boarding process as inputs to the model. For example, NPOs often can estimate that some volunteers may be more or less flexible in the types of tasks they are willing to work on, some volunteers may be less dedicated to participating, whereas other volunteers may have higher preference to work on particular tasks based on the information entered during on-boarding. Our method explicitly captures that there will be uncertainty in a NPO's estimate of volunteer task preferences and to capture this uncertainty we repeatably solve a deterministic model for different volunteer willingness scenarios.

Next, a consensus integer program is implemented to synthesize tasks offered on menus over multiple scenarios and delivers a final task recommendation menu that is personalized for each volunteer. The consensus menu is a final menu that is obtained by identifying tasks that appeared with most frequency over multiple scenarios of menus in the MC method.

From the menu offered to each volunteer, the volunteer's preference to work on each of the tasks in their menu is collected in terms of yes/no response (deemed as volunteer's willingness). Next, the volunteer group creation and task assignment model use the collected volunteer willingness values and only assigns volunteers tasks they are willing to work on. We formulate this group creation and assignment model as a deterministic integer program that uses volunteer demographic and skill information to create and assign ideal volunteer groups to tasks based on the learning potential and affinity levels among volunteers. The model's objective is to maximize the number of volunteers assigned to tasks while penalizing if assignments do not meet minimum and maximum group sizes, a task's skill level requirement, skill differences among volunteers within a group, or similar enough age IDs among volunteers in a group.

3.1 Model Assumptions and Notations

Our methodology is based on the following list of assumptions and notations.

Assumptions:

1. The NPO has a known list of "active" tasks and a known list of "active" volunteers who the NPO will send out personalized menus of tasks. Task skill level requirements required for successful completion of tasks and targeted minimum and maximum number of volunteers to be assigned per task, are pre-determined by the NPO when planning NPO projects/programs and designing tasks.
2. The NPO has access to some estimate of each volunteer's current skill-levels and task preferences, collected from volunteers during their onboarding process.
3. The NPO receives volunteer feedback on task preferences in the form of a yes or no response for tasks(s) they have been offered on their menus. If a volunteer is willing to work on a particular task, their 'yes' response is recorded as a value '1' and not selecting a task is recorded as a value '0'.
4. After a volunteer is assigned to a task, we assume that they work on the task till completion.
5. We assume that a homogenous volunteer group is an 'ideal' volunteer group. A homogeneous group in our methodology is a group where the volunteers assigned to a task in a group have similar age ID values and similar skill level values. We assume that similar age IDs among volunteers in a group will help create an ideal work environment for increased learning potential

and cohesion among volunteers in a group (Alam & Campbell, 2017; Baruch et al., 2016; Mazlan et al., 2018). while similar skill levels among volunteers in a group will support gradual, iterative, long-term, and successful learning is provided (Wiens et al., 2022).

6. The methodology assumes a group meets desired skill learning requirements if the group of volunteers consists of a set of volunteers that have at most a pre-defined difference in a given skill. The methodology assumes a volunteer is assigned a task to support collaboration, networking, and engagement if the volunteer is assigned to a task in a group with a volunteer that is of the same age ID or smaller than a predefined value.
7. A single-period static formulation is presented, where the NPO can use the methodology periodically (once per day/ once per week or every time a project manager needs to identify volunteer group to task assignments), yet decisions in one period are assumed to not influence the next time period.

Methodology Notations:

Sets:

V is the set of all volunteers, indexed on v

R is the set of all tasks, indexed on r

V' is the set of all mock volunteers, indexed on v'

R' is the set of all mock tasks, indexed on r' S is the set of all skills, indexed on s

Q is the set of all scenarios, indexed on σ

The sets of mock tasks R' is used to capture situations where volunteers do not have enough tasks for their ideal menu size to be met. The set of mock volunteers $V'$are used to capture situations where there are not enough tasks to meet ideal group sizes.

Parameters:

$\check{g}_r$= Maximum number of volunteers that can be assigned to task $r \in R$ (maximum allowed group size)

$\hat{g}_r$= Minimum number of volunteers that can be assigned to a task $r \in R$ (minimum group size)

b = Maximum number of tasks that can be recommended to a volunteer (maximum menu size)

d = Desired difference between skill levels of volunteers in a group to provide ideal learning environment (desired difference to be maximum d)

a = Desired difference between age IDs of volunteers in a group to obtain ideal affinity levels within groups (desired difference to be maximum a)

E_{rs} = the collective competency level of skill $s \in S$ required from a group of volunteers to successfully complete task $r \in R$ (value on the scale of 0-10 set by the NPO)

C_{vs} = current competency level of volunteer $v \in V$ on skill $s \in S$, given in integer values on the scale of 0-10

H_v = the age ID of volunteer $v \in V$ (ranging on the scale of 1 – 10) derived from the volunteer age collected as part of demographic information

n_1 = organization's preference weight for making volunteer to task assignments

n_2 = organization's preference weight for meeting task group size requirements

n_3 = organization's preference weight for meeting task skill requirements

n_4 = organization's preference weight for meeting group creation requirement that supports volunteer skill learning (by capturing difference in volunteer's skill levels)

n_5 = organization's preference weight for meeting group creation requirement that supports collaboration, networking, and engagement among volunteers (by capturing difference in volunteer's age IDs)

n_6 = organization's preference weight for meeting a menu size requirement

M = very large number

$\widehat{W}_{rv}$= 1 if the NPO expects volunteer $v \in V$ is willing to work on task $r \in R$; 0 otherwise

$\dot{W}_{rv}$= the actual willingness response collected from volunteers $v \in V$ to work on task $r \in R$, after being presented a final menu obtained as output from the MC IP (value 1 if volunteer v is willing to work on it and 0 otherwise)

The actual willingness response parameter ($\dot{W}_{rv}$) is a parameter used only in the GC IP model, while the expected willingness parameter ($\widehat{W}_{rv}$) is used only in the MC IP model. The remaining parameters are used in both the MC and GC IP models.

Decision Variables:

X_{rv} = 1 if task $r \in R$ is recommended to volunteer $v \in V$, 0 otherwise

$X'_{r'v}$= 1 if a mock task $r' \in R'$ is recommended to volunteer $v \in V$, 0 otherwise. This binary variable is used to capture and penalize in the objective function when there are not enough tasks $r \in R$ to be presented to a volunteer on their menu to meet the desired menu size b

Z_{rv}= 1 if task $r \in R$ is assigned to volunteer $v \in V$, 0 otherwise

$Z'_{rv'}$= 1 if task $r \in R$ is assigned to mock volunteer $v' \in V'$, 0 otherwise. This binary variable is used to capture and penalize in the objective function when volunteers $v \in V$ cannot meet the pre-defined minimum ($\hat{g}_r$) or maximum ($\check{g}_r$) group sizes when task assignments are made

F_{rs} = positive integer value by which the group of volunteers assigned to a task does not meet the collective skill level (E_{rs}) in skill $s \in S$ required on task $r \in R$

β_{rs} = 1 if value of F_{rs} is greater than 0, 0 otherwise. This binary variable is used to capture instances where the collective skill level required in skill $s \in S$ on a task $r \in R$ was not met by the group of volunteers assigned to the task

D_{rs} = positive integer value by which the difference in skill level in skill $s \in S$ on task $r \in R$ among volunteers assigned in a group does not meet the pre-defined value d

λ_{rs} = 1 if value of D_{rs} is greater than 0, 0 otherwise. This binary variable is used to capture the instances where the difference in skill level in skill $s \in S$ on task $r \in R$ does not meet the pre-defined value d

A_v = positive integer value by which the difference in age ID value among volunteers $v \in V$ assigned in a group does not meet the pre-defined value a

ρ_v = 1 if the value of A_v is greater than 0, 0 otherwise. This binary variable is used to capture instances where the age ID difference among any two volunteers $v \in V$ assigned in a group does not meet the pre-defined value a

G_{rs} = maximum value of skill competency in skill $s \in S$ among the volunteers assigned to task $r \in R$

I_{rvs} = 1 if on task $r \in R$, volunteer $v \in V$ is identified as possessing the maximum skill competency in skill $s \in S$, 0 otherwise

J^1_{rvs} = 0 if $Z_{rv} = 1$; 0 otherwise. Implemented to ensure constraints (16) – (18) capture the values of skill difference in skill $s \in S$ if a volunteer $v \in V$ has been assigned to a task $r \in R$.

$J^2_{rv_1v_2}$ = 0 if $Z_{rv_1} + Z_{rv_2} = 2$; 1 otherwise. Implemented to ensure constraints (23) – (25) are applied only when a group of at least two or more volunteers $v \in V$ have been assigned to a task $r \in R$

In our model, the variable X_{rv} captures which task $r \in R$ is recommended to volunteer $v \in V$, and there are multiple task(s) r that can be offered on a volunteer v task recommendation menu. The variable Z_{rv} captures the assignment of task $r \in R$ to volunteer $v \in V$. These two variables are used so that a volunteer can be offered more than one task per task recommendation menu offered, but a volunteer can be assigned to at most one task.

3.2 Menu Creation Integer Program for a Given Volunteer Task Willingness Scenario

This section presents a deterministic integer linear programming model in equations (1) – (31) that takes inputs of volunteer and organizational information and outputs menus of task recommendations personalized for each volunteer. The Menu Creation (MC) integer program is run multiple times with different data representing different volunteers' varying task preferences. We denote one such input run as a scenario σ ∈ Q.

Maximize $n_1 \sum_{v \in V} \sum_{r \in R} Z_{rv} - (n_2 \sum_{v' \in V'} \sum_{r \in R} Z'_{rv'} + n_3 \sum_{r \in R} \sum_{s \in S} \beta_{rs} + n_4 \sum_{r \in R} \sum_{s \in S} \lambda_{rs} + n_5 \sum_{v1 \in V} \rho_v + n_6 \sum_{v \in V} \sum_{r' \in R'} X'_{r'v})$ (1)

Subject to:

$\sum_{r \in R} X_{rv} + \sum_{r' \in R'} X'_{r'v} = b \qquad \forall v \in V \qquad (2)$

$Z_{rv} \le X_{rv} \qquad \forall r \in R, \forall v \in V \qquad (3)$

$X_{rv} \le \widehat{W}_{rv} \qquad \forall r \in R, \forall v \in V \qquad (4)$

$\sum_{r \in R} Z_{rv} \le 1 \qquad \forall v \in V \qquad (5)$

$$\sum_{v \in V} Z_{rv} + \sum_{v' \in V'} Z'_{rv'} \leq \check{g}_r \qquad \forall r \in R \qquad (6)$$

$$\sum_{v \in V} Z_{rv} + \sum_{v' \in V'} Z'_{rv'} \geq \hat{g}_r \qquad \forall r \in R \qquad (7)$$

$$\sum_{v \in V} Z_{rv} \ C_{vs} \geq E_{rs} - F_{rs} \qquad \forall r \in R, \forall s \in S \qquad (8)$$

$$F_{rs} \leq M\ (\beta_{rs}) \qquad \forall r \in R, \forall s \in S \qquad (9)$$

$$G_{rs} \leq (Z_{rv} C_{vs}) + M(1 - I_{rvs}) \qquad \forall v \in V, \forall r \in R, s \in S; \text{where } E_{rs} \geq 1 \qquad (10)$$

$$\sum_{v \in V} I_{rvs} \geq 1 \qquad \forall r \in R, \forall s \in S; \text{where } E_{rs} \geq 1 \qquad (11)$$

$$G_{rs} - Z_{rv} \ C_{vs} \geq (1 - I_{rvs}) \qquad \forall v \in V, \forall r \in R, s \in S; \text{where } E_{rs} \geq 1 \qquad (12)$$

$$G_{rs} - (Z_{rv} \ C_{vs}) \leq d + D_{rs} + (M\ J^1_{rvs}) \qquad \forall v \in V, \forall r \in R, \forall s \in S \qquad (13)$$

$$Z_{rv} \leq 1 + M(1 - J^1_{rvs}) \qquad \forall v \in V, \forall r \in R, \forall s \in S \qquad (14)$$

$$Z_{rv} \geq 1 - M\ J^1_{rvs} \qquad \forall v \in V, \forall r \in R, \forall s \in S \qquad (15)$$

$$D_{rs} \leq M\ (\lambda_{rs}) \qquad \forall r \in R, \forall s \in S \qquad (16)$$

$$\left(Z_{rv_1} H_{v_1} - Z_{rv_2} H_{v_2}\right) \leq a + A_{v_1} + (M\ J^2_{rv_1v_2}) \qquad \forall v_1, v_2 \in V; \text{where } v_1 \neq v_2, \forall r \in R \qquad (17)$$

$$\left(Z_{rv_1} + Z_{rv_2}\right) \leq 2 + M(1 - J^2_{rv_1v_2}) \qquad \forall v_1, v_2 \in V; \text{ where } v_1 \neq v_2, \forall r \in R \qquad (18)$$

$$\left(Z_{rv_1} + Z_{rv_2}\right) \geq 2 - M\ J^2_{rv_1v_2} \qquad \forall v_1, v_2 \in V; \text{ where } v_1 \neq v_2, \forall r \in R \qquad (19)$$

$$A_v \leq M\ (\rho_v) \qquad \forall v \in V \qquad (20)$$

$$X_{rv}, X'_{r'v}, Z_{rv}, Z'_{rv'}, \beta_{rs}, \lambda_{rs}, \rho_v, I_{rvs}, J^1_{rvs}, J^2_{rv_1v_2} \in \{0,1\}$$
$$\forall r \in R, \forall v \in V, \forall s \in S, r' \in R, \forall v' \in V' \qquad (21)$$

$$A_v, D_{rs}, F_{rs}, G_{rs} \in \mathbb{Z}^+ \qquad \forall r \in R, \forall v \in V, \forall s \in S \qquad (22)$$

The objective function in equation (1) maximizes the number of volunteer-to-task assignments (Z_{rv}) made while subtracting a weighted count of penalties if assignments do not meet any of the following conditions: minimum and maximum group sizes ($Z'_{rv'}$), a task's skill level requirement (β_{rs}), the skill difference among volunteers within a group (λ_{rs}), the age ID difference among volunteers in a group (ρ_v), and the menu size ($X'_{r'v}$). Constraint (2) enforces a menu size for the task menu recommended to a volunteer. The mock task recommendation variables are implemented to prevent infeasibility by recommending mock tasks r' on menus presented to volunteers when enough tasks r are not available, with the count of mock tasks being penalized in the objective function. Constraint (3) ensures that a volunteer v can only be assigned to a task if they were offered the task on their menu recommendation. Constraint (4) ensures that the tasks offered to a volunteer in their personalized menu recommendation are tasks that the NPO expects them to be willing to work on. Constraints (5)-(20) are constraints developed for the Group Creation IP (Section 3.4) and are implemented in the menu creation IP to recommend task menus to volunteers that have the potential to create ideal groups assignments. Constraint (5) ensures that a volunteer is assigned to at most one task. Constraints (6)-(7) enforce maximum and minimum group sizes for groups created for task assignment. For instances where a targeted maximum or minimum group size is infeasible (e.g., insufficient number of volunteers), a mock volunteer assignment decision variable ($Z'_{rv'}$) allows for volunteer to task assignments to still be made and invokes a penalty in the objective function. Constraint (8) ensures that the collective skill competency level of all volunteers assigned to a task meets the task's pre-defined skill level requirement value (E_{rs}), and variable F_{rs} captures the value by which a volunteer group assignment does not meet E_{rs}. Constraint (9) is implemented to obtain a count of all skills(s) $s \in S$ on task(s) $r \in R$ on which the collective skill level requirement value (E_{rs}) was not met, thus invoking a penalty (β_{rs}) in the objective function for all instances where skill requirements are not met for every task-skill combination. Constraints (10)-(12) identify the maximum skilled volunteer within a group assigned to a task (I_{rvs}), and this is utilized to create groups with at least one volunteer with a higher skill level than all others assigned in a group to ensure skill learning occurs within created groups. Constraints (13)-(15) ensure the difference in skill levels among the highest skilled volunteer and other volunteer(s) assigned in a group to a task is less than or equal to a pre-set, value d. In instances where the difference in skill values among volunteers is greater than d, variable D_{rs} captures the value by which volunteers assigned in a group do not meet value d, thus invoking a penalty (λ_{rs}) in the objective function. Constraint (16) is implemented to obtain a count of instances where the value for variable D_{rs} is a positive integer value, thus giving a count of instances where the skill difference among volunteers is not at a desired value in created groups. Constraints (13)-(16) are implemented to create homogeneous volunteer groups based on skill levels to improve learning potential and skill transfer among volunteers working together. Constraints (17)-(19) ensure that the difference in age ID values among volunteers assigned in a group to a task is less than or equal to a pre-set, 'age ID difference desired value' a. In instances where the difference in age ID values among volunteers is greater than a, variable A_v captures the value by which the age ID value among volunteers assigned in a group does not meet value a, thus invoking a penalty (ρ_v) in the objective function. Constraint (20) is implemented to obtain a count of instances where variable A_v is a positive integer value, thus giving a count of instances where age ID requirement is not met in group creations. Constraint (21) enforces binary decision variables that can only get assigned values of 0 or 1. Constraint (22) indicates the variables that are positive integer values.

The Menu Creation IP when run for different volunteer willingness scenarios σ ∈ Q, produces an output of different menus for each scenario σ, personalized to every volunteer $v \in V$. The Menu Creation IP is solved separately and repeated for each scenario σ ∈ Q, and the resulting task recommendation menus X_{rv} for these different scenarios are used as an input into the Consensus Integer program in section 3.3.

3.3 Consensus Integer Program

To identify which task(s) should be recommended to each volunteer on the final menu presented to them, the consensus IP is solved to identify the task(s) that appear most frequently across menus produced for all scenarios. The consensus IP consists of the decision variable X_{rv}, which captures the tasks to be offered on the final menu, per volunteer. The number of times a task $r \in R$ is recommended to volunteers $v \in V$ is an input parameter to the Consensus IP. This is defined as the frequency $Freq_{rv}$, which is the frequency (total number of times) the task $r \in R$ is recommended to volunteer $v \in V$ summed over all the scenarios. The Consensus Integer Program is solved with objective function Maximize $\sum_{v \in V} \sum_{r \in R} X_{rv} \; Freq_{rv}$, subject to a menu size constraint: $\sum_{r \in R} X_{rv} \leq b \quad \forall \, v \in V$.

The final task recommendation menus (X_{rv}) obtained from the Consensus IP is presented to each volunteer. Each volunteer then provides their willingness to work on any of the particular task(s) presented to them on their menu. This volunteer willingness information is then used as an input into the group creation and task assignment model ($\dot{W}_{rv} = 1$ if volunteer v is willing to volunteer on task r, 0 otherwise). This information, along with an estimate of volunteer's demographic and skill information, is used as inputs for the group creation and task assignment IP.

3.4 Group Creation and Task Assignment Integer Program

The Group Creation and Task Assignment IP maximizes (23) subject to constraints (5)-(20), and (24)-(26), which creates volunteer groups and their assignments to NPO tasks.

$$\text{Maximize } n_1 \sum_{v\in V} \sum_{r\in R} Z_{rv} - (n_2 \sum_{v'\in V'} \sum_{r\in R} Z'_{rv'} + n_3 \sum_{r\in R} \sum_{s\in S} \beta_{rs} + n_4 \sum_{r\in R} \sum_{s\in S} \lambda_{rs} + n_5 \sum_{v\in V} \rho_v) \quad (23)$$

Subject to:

$$Z_{rv} \leq \dot{W}_{rv} \qquad \forall r \in R,\ \forall v \in V \qquad (24)$$

$$Z_{rv}, Z'_{rv'}, \beta_{rs}, \lambda_{rs}, \rho_v, I_{rvs}, J^1_{rvs}, J^2_{rv_1v_2} \in \{0,1\} \qquad \forall r \in R,\ \forall v \in V, \forall s \in S \qquad (25)$$

$$A'_v, D'_{rs}, F'_{rs}, G_{rs} \in \mathbb{Z}^+ \qquad \forall r \in R,\ \forall v \in V, \forall s \in S \qquad (26)$$

The objective function in equation (23) maximizes the number of volunteer-to-task assignments made while subtracting the weighted count of penalties if assignments do not meet any of the following conditions: minimum and maximum group sizes ($Z'_{rv'}$), a task's skill level requirement (β_{rs}), the skill difference among volunteers within a group (λ_{rs}), and the age ID difference among volunteers in a group (ρ_v). In this model, we also implement group creation constraints that have been explained in the menu creation model constraints (5)-(20) (see Section 3.2). Constraint (24) enforces the assignment of a volunteer to a task only if they have responded back with a positive willingness to work on the respective task ($\dot{W}_{rv}$), which is collected from the volunteers after presenting their personalized menus, created using the menu creation IP and the consensus IP. Constraints (25) enforce the binary decision variables that can only get assigned values of 0 or 1. Constraint (26) indicates the variables that are positive integer values.

4. NPO Empirical Data Analysis

In this section, we present the empirical data analysis carried out on a volunteer data set obtained in 2023 from a partner NPO who utilizes remote volunteers to work on online content creation tasks. We summarize the NPO's volunteer information dataset at an aggregate level to obtain information that we use to generate representative data for our experiments (see Appendix X1,X2).

The NPO has remote volunteers participating from various countries across the world working on online content creation tasks. In the NPO's current process flow, volunteers sign-up on the NPO's website, where the NPO collects preliminary volunteer information (volunteer name, email address, city, country, age, fluent language(s), career/educational qualifications, past volunteering experience, availability, task preferences (up-to 5 choices of task selections from a pre-populated list), and new project ideas). Volunteers are currently assigned to tasks manually by senior volunteers or employees. While the NPO does not explicitly capture skill requirements for tasks, their experienced senior volunteers and staff use their understanding of what each task requires and utilizes volunteer's responses on education/career, past volunteering experience and language abilities when they make volunteer assignments. The NPO's tasks do not need to be completed during specific time periods or shifts and can be completed by the volunteer(s) in their own time as long as they are completed by a pre-defined due date, which is typically in a few weeks or months (tasks volunteers help with are usually not urgent). The NPO has twelve types of online tasks, which we assigned unique task IDs (1-12): 1-Data Entry, 2-Blog Writing, 3-Animation, 4-App Development, 5-Graphic Design, 6-Subtitle Writing, 7-Translation, 8-Web Development, 9-Video Editing, 10-Pamphlet Creation, 11-Social Media Management, 12-Audio Content Manager. Volunteers are needed for each task, to work towards the completion of a particular project initiated by the NPO. For example, a 'Kiddie Group' project that requires creating content for a young age group, will consist of tasks such as animation, video editing, translation, subtitle writing, and graphic design.

First, we review volunteer ages from the volunteer dataset. We define a set of age ranges and associated age ID values (Age IDs 1-10) as follows: Age ID 1 for volunteers of ages below 15, 2 for 15-19, 3 for 20-24, 4 for 25-29, 5 for 30-34, 6 for 35-39, 7 for 40-44, 8 for 45-49, 9 for 50-54, and 10 for volunteer ages between 55-60. We then identify the distribution of volunteers in the NPO's dataset for each age ID, that selected online tasks. According to the distribution presented in Figure 3, a large majority of the volunteer population with the NPO's online tasks is young, and no volunteers above the age of 45 participated with the NPO in online tasks. The input parameter H_v indicates the age ID of volunteer $v \in V$, and we generate data for our computational experiments using the discrete probability distribution of volunteers in each age ID category (1-10) shown in Figure 3.

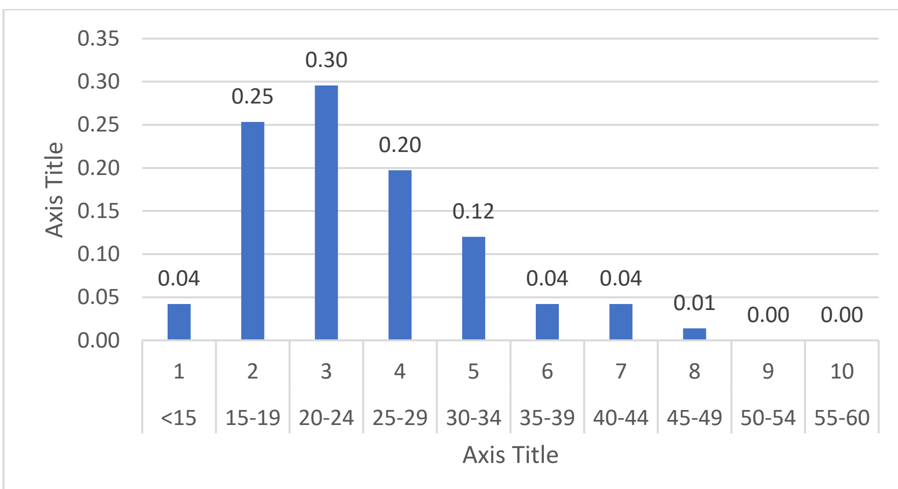
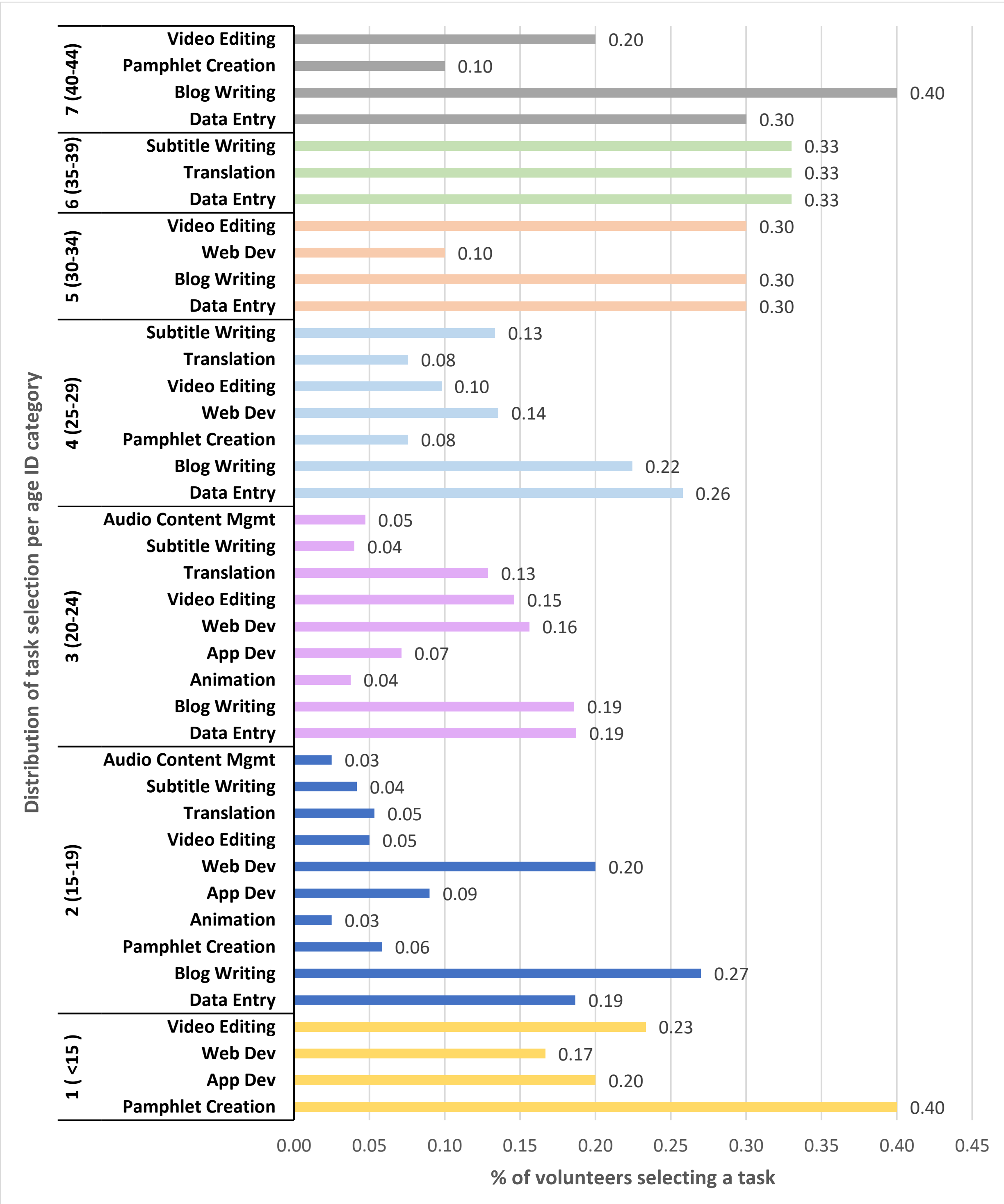


Figure 3: The probability distribution of volunteers that selected to work on online tasks, in each age ID category

It is observed that the distribution of task selections made by volunteers (in the NPO's initial onboarding form) varies per age ID, indicating that a volunteer's preference for working on a particular task is influenced by their age ID. Figure 4 displays the probability distributions of a volunteer's task selections (x-axis), conditioned on the volunteer belonging to a specific age ID category, for all NPO tasks (y-axis). For every age ID category displayed (1-7), only the tasks with some positive probability of being picked by volunteers of that age ID category are displayed. Tasks with zero probability of being selected for a given age ID are not presented. Volunteers of age ID 8-10 in the NPO's dataset did not select any online tasks on their onboarding form, and only select the in-person tasks offered by the NPO, and thus, we only display task selection of volunteers belonging to the first seven age ID categories.



*** Tasks:** Data Entry, Blog Writing, Pamphlet Creation, Animation, App Development, Web Development, Video Editing, Translation, Subtitle Writing, Graphic Design, Social Media Mgmt, Audio Content Mgmt

**** Age IDs:** 1 - 7

Figure 4: The conditional probability distributions of task selections for different volunteer age ID categories

As observed in Figure 4, volunteers of the older age ID categories tend to pick tasks that are simple to complete or tasks that they have the knowledge of through education or experience. For example, volunteers above the age of 30 tended to select either simple tasks e.g., Data Entry, Translation, Subtitle Writing, Blog Writing, Pamphlet Creation, or high intensity tasks they are skilled in through education/work e.g., Video Editing, Web Development. Figure 4 also shows that the younger volunteers (age ID 2+) tend to select a wider range of tasks they are interested in volunteering, likely influenced by their desire to gain experience through the tasks they complete. The data table of the conditional probability distributions of volunteer's task selections per age ID category is presented in Appendix X5.

Next, we summarize volunteer task priority selection information. Task priority selection is information collected by the NPO using their initial volunteer onboarding form, and refers to task preference(s) selected by a volunteer As shown in Figure 5, the total number of task priority selections that a volunteer makes on their onboarding form varies, with the majority (54% selecting 5 tasks), yet many select fewer, with even some (8% of volunteers) only selecting one task. As described in our volunteer choice model in Section 5.1, we associate the number of task selection(s) made by a volunteer on their onboarding form as an identifier of their eagerness or dedication to select and work on a task presented to them: volunteers selecting 5 task preferences are considered more dedicated than volunteers selecting only 1 task preference.

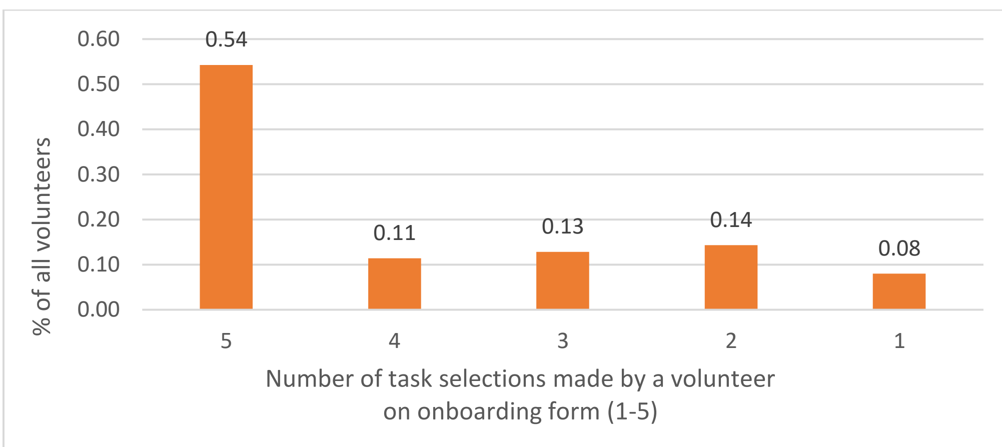


Figure 5: The probability distribution that a volunteer selected a given number of tasks during onboarding

Next, with an understanding of the NPO tasks, we define the following 10 skills: S1-coding and development, S2-basic computer skills, S3-languages, S4-writing, S5-animation, S6-video editing, S7-language fluency, S8-graphic design, S9-social media, S10-analytics. We convert the volunteer responses to open-ended questions asked during volunteer onboarding into discrete probability distributions denoting a volunteer's skill levels for the defined set of skills. The open-ended data fields available are 'What languages are you most fluent in', 'What other languages are you strongly familiar with', 'Current career/educational qualifications', 'Past volunteering experiences', and 'New project ideas', where volunteers' open responses consist of terms that indicate their skill(s) (see Appendix X1 and X2). For example, a volunteer's skill level for 'coding and development' is identified from the data fields 'Current career/educational qualifications' and 'Past volunteering experiences' where volunteer's responses are checked for terms 'software', 'computer', 'engineering', 'code', and 'web', etc. For each volunteer in the database, we then obtain a count of terms extracted to obtain skill levels of volunteers and these counts are then used to identify the distribution of low (1), to high (7) skilled volunteers for each specific skill. Figure 6 shows the resulting probability distribution of volunteers with skill levels 1 (low)-7 (high) for skills 1 – 10.

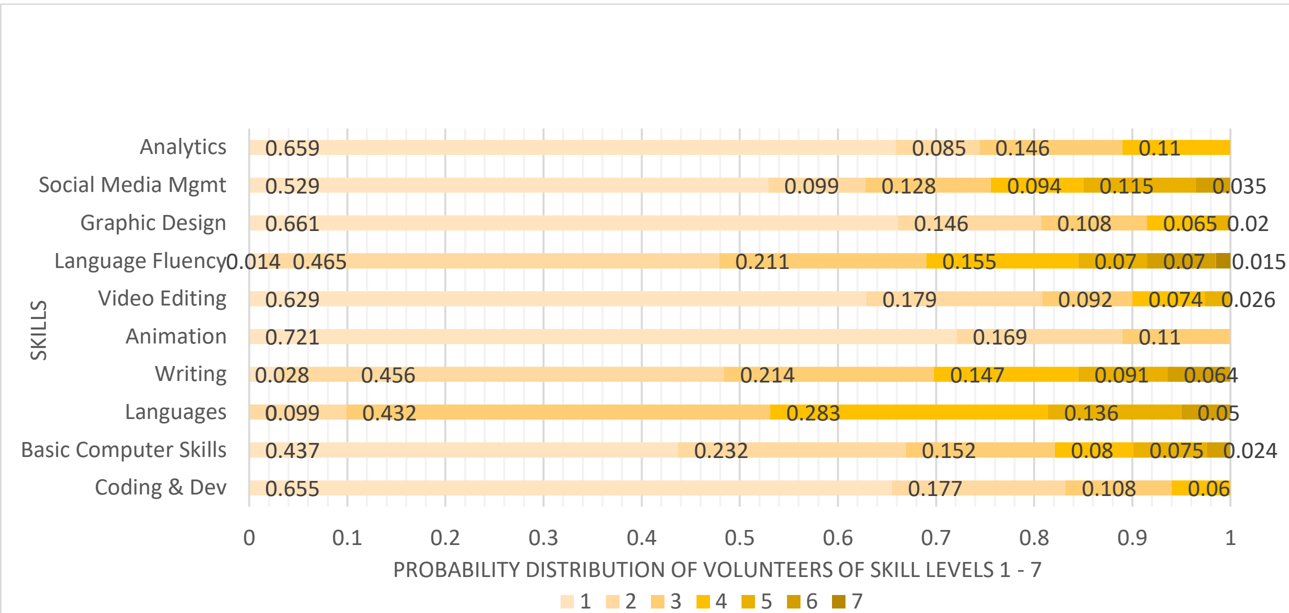


Figure 6: The probability distribution of volunteers with different levels of skills

In Figure 6, we observe that the volunteer population being analyzed is highly skilled in language related skills (e.g., S3-language and S4-writing), and less skilled in technology related skills (e.g., S5-animation, S6-video editing, etc.). The input parameter C_{vs} is the value of a volunteer's skill level $v \in V$ in skill $s \in S$, with values between 1-7, and in our computational experiments we generate these values using the distribution of volunteer skills of volunteers working with the NPO (Figure 6). The data table of the probability distribution of volunteers with different skill levels (1-7) for the 10 skills we have defined from the NPO's volunteer dataset (Appendix X2) is presented in Appendix X6.

5. Computational Experimental Set-up

We design our computational experiments to help identify the impact of our methodology on providing volunteers with autonomy in task selection and the impact of considering ideal groups when making volunteer to task assignments in groups. We focus on answering the following questions: (1) what is the impact of implementing vs. not implementing the Menu Creation methodology to a NPO? (2) what is the impact of environmental factors (picky volunteers, heterogeneity in task selections) on successfully achieving ideal volunteer groups? (3) what is the impact of implementing learning potential and affinity constraints for volunteer group creation and on meeting NPOs task needs? (4) what is the impact of implementing a minimum group size ($\hat{g}_r$) when creating volunteer groups? To answer these questions, we design computational experiments and generate synthetic but representative volunteer and NPO data for experiments using the empirical data from our NPO partner described in Section 4. A detailed description of the generation of our representative data can be found in Appendix X8. In Section 5.1 we describe the volunteer choice model used to model volunteer willingness behavior. In Section 5.2 present a summary of the factors and their levels, as well as the benchmark policies designed and implemented to analyze the impact of our methodology. In Section 5.3 we define the various key performance indicators used to evaluate the impact of our methodologies.

5.1 Volunteer Choice Model

In our experiments, we estimate a volunteer's preference for a task using the multi-attribute utility function given in (27). This function captures that the NPO expects volunteer v's preference for task r in scenario σ (which we denote as α_{rv}^{σ}) is made up of two attributes (i_v and j_{rv}) and an external factor that captures any deviation not associated with the two attributes.

$$\alpha_{rv}^{\sigma} = i_v * 2 * \delta_v^{\sigma} + j_{rv} * 10 * (1 - \delta_v^{\sigma}) + \epsilon_{rv}^{\sigma} \qquad \forall r \in R, \forall v \in V \quad (27)$$

The attributes used in α_{rv}^{σ} are chosen based on task attributes found to impact volunteer's likelihood to select and work on a task and are established from volunteer information available to the NPO, and qualitative studies (Bang, 2015; Harrison, 1995; Millette & Gagné, 2008). The input parameter j_{rv} is the probability of a volunteer selecting a task $r \in R$, given the age ID of the volunteer is H_v, with values between 0-1. In our computational experiments, we generate these using conditional probability distributions of task selections made by volunteers belonging to each age ID category, obtained from our partner NPO's volunteer dataset (see Figure 4). The input parameter i_v is the number of task preference selections made by volunteer $v \in V$ in their onboarding form entry and in our computational experiments, we generate these values using the discrete probability distribution of the number of volunteer's task preference selections obtained from our partner NPO's volunteer dataset (see Figure 5). δ_v^{σ} is volunteer's $v \in V$ attribute weight that represents the weight put on attributes i_v versus j_{rv}, and varies per scenario σ. In our experiments we generate δ_v^{σ} using a uniform distribution with values between a minimum of 0 and maximum of 1. Finally, the variable ϵ_{rv}^{σ} captures a volunteer's utility from sources other than the two attributes i_v and j_{rv}. In our experiments we randomly generated these using a Gumbel distribution with a mean of 0 and standard deviation of 1, for all $v \in V$, $r \in R$, and σ $\in Q$.

We also estimate a volunteer's minimum expected preference value that a volunteer v is willing to volunteer for, in scenario σ, which we denote as α_{0v}^{σ}. We generate α_{0v}^{σ} using equation (28), where two terms, ń and ϵ_v are added to generate values for four different environmental instances. The term ń is generated using a normal distribution, as it allows us to generate values to represent, allowing us to represent four different volunteer populations in our experiments. Mean values of 3 and 8 are used to represent low and high pickiness levels of volunteers. A standard deviation value of 0 represents a set of volunteers displaying no heterogeneity in their pickiness behaviors and a value of 1 represents a set of volunteers displaying some heterogeneity in their pickiness behaviors. An error term ϵ_v is generated with a Gumbel distribution with mean 0 and standard deviation 1.

$$\alpha_{0v}^{\sigma} = \acute{n} + \epsilon_v \qquad \forall\, v \in V \qquad (28)$$

In the Menu Creation model presented in section 3.2, if the utility value α_{rv}^{σ} for volunteer $v \in V$ and task $r \in R$ is greater than the utility threshold value α_{0v}^{σ}, the expected willingness value for the volunteer v to work on the task r, $\widehat{W}_{rv}$ is set at 1 in scenario σ, and the NPO expects the volunteer to select the task if presented to them, and set at zero otherwise.

After the volunteers have been presented with their personalized final task recommendation menus, we generate $\dot{W}_{rv}$, which is the actual willingness of volunteer $v \in V$ to work on task $r \in R$. $\dot{W}_{rv}$ is generated as a 1 if their actual utility value for working on a task α_{rv} (sampled using equation (36)) is greater than their minimum actual preference value α_{0v} (sampled using equation (28)). $\dot{W}_{rv}$ set to 1 indicates that a volunteer has stated their willingness to work on a task presented to them and is set to 0 otherwise.

5.2 Factors and Levels for our Design of Experiments:

In our experiments, we vary five factors summarized in Table 1, of which the volunteer pickiness values (via the volunteer's willingness thresholds α_{0v}), is an exogenous factor, which we assume is outside of the NPO's control, and represents four different operating environments where volunteers have different pickiness for taking on tasks presented to them, and display different pickiness variation among the set of active volunteers. To identify the impact of implementing key features of our methodology, we capture implementing versus not implementing particular factors. Specifically, the benchmark experiments (Table 1 - in bold) are used to quantify the impact of our method's components, i.e., offering volunteers with task recommendation menus, implementing group creation constraints (skill level difference and age ID difference among group members) to obtain 'ideal' groups, and implementing a constraint for minimum group sizes allowed. Using a full factorial design, results are collected for 96 different instance settings that are unique combinations of all factor levels of the design of experiments.

For each experiment run, we run 5 replications (where volunteer input parameter data generated changes per replication), with 10 scenarios per replication (to model different task selection preferences). The number of replications and scenarios to be run for experiments is decided based on results observed in Appendix X3.

The values for the number of tasks, volunteers and the maximum group size allowed are fixed in our experiments as following:

- Set of tasks $|R|$ = 24
- Set of volunteers $|V|$= 72
- Max group size $\check{g}_r$= 3

The value for the maximum group size is chosen as findings on group creation in qualitative literature note that larger groups are less successful at supporting learning potential and collaboration due to too many different individuals within one group (Barnes & Sharpe, 2009).

Factor	Levels
Min group size $\hat{g}_r$	**1 (no groups enforced)**
	2
Max menu size b	**1 (no menus formed)**
	3
	5

Maximum desired skill difference among volunteers in groups (d)	**7 (no skill difference enforced among volunteers in a group)**
	1
Maximum desired age ID difference among volunteers in a group (a)	**7 (no age ID difference enforced among volunteers)**
	1
α_{0v} generation – Equation (28)	ń $\sim N$(mean =3, SD = 0) (Low picky volunteers, no heterogeneity)
	ń $\sim N$(mean =8, SD = 0) (High picky volunteers, no heterogeneity)
	ń $\sim N$(mean =3, SD = 1) (Low picky volunteers, some heterogeneity)
	ń $\sim N$(mean =8, SD = 1) (High picky volunteers, some heterogeneity)

Table 1: A full-factorial design with exogenous factors and levels

5.3 Stakeholder Key Performance Indicators:

We aim to quantify how our developed methodology makes task assignments that encourage volunteers' varying motivations and helps the NPO fulfill high levels of task assignments with volunteers. To do so, we developed and collected values for the following KPIs:

(i) - Objective function value output obtained from the GC IP model which is calculated as the value of maximum number of volunteer to task assignments, minus penalties if assignments made were not able to meet: set minimum and maximum group sizes ($Z'^{\sigma}_{rv'}$), a task's skill level requirement (β_{rs}), the value of skill difference among volunteers within a group (λ_{rs}), the value of age ID difference among volunteers in a group (ρ_v) (Section 3.4, equation (23)).

(ii) - Total number of volunteers assigned which is captured as the total number of volunteer to task assignments made.

(iii) - Number of tasks with volunteer(s) assigned, which is captured as the number of tasks that were either assigned a single volunteer or a group of volunteers.

(iv) - Number of volunteers assigned in groups, which is captured by calculating the number of volunteers assigned to tasks in groups of size 2 of greater = Total number of volunteers assigned – Number of single volunteer assignments to tasks.

(v) - % of volunteer pairs with similar age IDs, which captures the number of volunteer pairs assigned together that have similar age ID values divided by the total number of groups formed and assigned to tasks.

(vi) - % of volunteers assigned in groups with similar skill level values, which is calculated as the number of volunteers assigned in groups (of 2 or more) that have similar skill level values, divided by the total number of groups formed and assigned to tasks.

(vii) - % of total task-skill requirements met, which is calculated as the number of task on which an assigned volunteer group collectively meets the skill levels required for all skills required on the task, divided by the total number of task-skill combinations on present on all projects of the NPO.

In addition to the results and analysis presented in the next section, we also ran ANOVA tests which illustrated that the KPIs were impacted by changing factor levels implemented in our design of experiments (see Appendix X7).

6. Results

In Tables 2, 3, 4 and 5, the following KPIs are presented: (i) – Objective function value[1], (ii) – Total number of volunteers assigned, (iii) – Total number of tasks with volunteer(s) assigned (of total 24 tasks), (iv) – Number of volunteers assigned in groups, (v) – % of volunteer pairs with similar age IDs, (vi) – % of volunteers assigned in groups with similar skill level values (vii) – % of total task-skill requirements met. Each of these tables present results broken down by their factor levels as previously defined for our design of experiments (see Table 1).

Table 2 presents a summary of results for all KPIs collected for every factor and level implemented on our design of experiments. The 'Average Over All Instances' row in Table 2 represents the average of values captured across all experiments for each KPI. This is followed by the average values conditioned on that they have a specific factor level (e.g., the average objective function value across all instances is 67.33, whereas the average objective function value for all instances with a minimum group size value set at 1 is 69.99).

Factor	Level	KPIs						
		(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Avg Over All Instances		67.33	45.89	19.54	42.17	0.79	0.62	0.71
Min Group Size Not Implemented (1), Implemented (2)								
Min Grp Size ($\hat{g}_r$)	**1**	69.99	45.88	19.62	41.82	0.80	0.62	0.71
	2	64.67	45.90	19.46	42.52	0.79	0.62	0.70
No Menu Created (1), Menu Created (Size 3 & 5)								
Max Menu Size (b)	**1**	55.80	37.64	17.81	32.54	0.78	0.57	0.59
	3	71.07	48.42	20.12	45.10	0.79	0.63	0.75
	5	75.12	51.65	20.69	48.86	0.82	0.66	0.79
Ideal Group Creation Cts Not Implemented (7,7), Implemented (1,1)								
Min Skill Diff (d) & Age ID Diff (a)	d **= 7,** a **= 7**	71.61	48.68	19.50	45.65	0.67	0.57	0.72
	d **= 1,** a **= 1**	63.30	42.82	19.63	38.28	0.92	0.66	0.70
	d **= 7,** a **= 1**	69.25	47.05	19.47	43.84	0.90	0.56	0.72
	d **= 1,** a **= 7**	65.15	44.93	19.56	40.96	0.69	0.67	0.69
Environments - Low Picky, High Picky, No Heterogeneity, Some Heterogeneity								
α_{0v} Generation	**LPV*, NH***	83.26	59.14	23.28	56.57	0.77	0.74	0.87
	LPV*, SH*	84.38	59.06	23.23	56.83	0.77	0.75	0.87
	HPV*, NH*	22.86	12.17	8.93	5.78	0.86	0.31	0.28
	HPV*, SH*	78.80	53.86	22.72	50.21	0.78	0.68	0.83

***** LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

Table 2: Summary of Results for all KPIs (a) – (g), summarized for every factor and level implemented in DOE

6.1 Impact of the Menu Creation Methodology

[1] The KPI (a) - Objective function value has been re-scaled from a range of negative and positive values to a range of values between 0 – 100 for a better representation of percent change values throughout the results section. Formula used to re-scale values presented in Appendix X4.

In this section, we quantify the impact of using the Menu Creation (MC) methodology. The benchmark policy for this factor is Menu Size b = 1, (i.e., when task recommendation menus are not created and only one task is presented to volunteers). In Table 3, we present results where we compare this benchmark policy (b = 1) to the MC methodology with a menu size of b = 3, and where we compare the benchmark policy (b = 1) to the MC methodology with a menu size of b = 5. Results for implementing the MC methodology with menu size of b = 3 vs. menu size of b = 5 being implemented are presented in Appendix X9. Results are presented for each of the four defined environmental factors (HPV, LPV, NH, SH).

Table 3 presents the % change between different implemented factor levels on KPIs (i) – (vii), where the % change is calculated using formula:

$$\% \text{ change } = \frac{(Value\ when\ MC\ implemented - Value\ when\ MC\ not\ implemented)}{(Value\ when\ MC\ not\ implemented)} * 100$$

We also conduct two-tailed t-tests and present the two-tailed p-value for the % change over all environmental instances for each KPI. In Tables 3, 4 and 5, p-values smaller than 0.05 indicate a significant impact on a given KPI due to the implementation of the menu creation methodology.

Avg. Change (%)	**Environmental Factors**	**KPIs**						
		(i)	**(ii)**	**(iii)**	**(iv)**	**(v)**	**(vi)**	**(vii)**
No Menu Created (b = 1) vs. Menu Size b = 3 Created								
b = 1 vs. 3	**Avg Over All Env. Instances**	27.36	28.64	12.99	38.45	1.03	10.71	28.54
	p-value	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05
	LPV*	20.60	22.80	6.57	31.67	4.34	5.56	20.98
	HPV*	40.14	40.28	23.71	54.16	-2.37	19.49	42.16
	NH*	30.26	32.21	16.98	41.63	-0.50	14.74	30.26
	SH*	25.53	26.51	10.36	36.68	2.36	7.93	27.90
No Menu Created (b = 1) vs. Menu Size b = 5 Created								
b = 1 vs. 5	**Avg Over All Env. Instances**	34.62	37.45	16.22	50.16	5.23	15.84	34.53
	p-value	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05	<<0.05
	LPV*	23.92	29.31	6.68	40.31	7.05	6.99	24.86
	HPV*	54.86	55.67	32.15	75.93	3.06	32.52	54.56
	NH*	38.33	40.08	24.76	49.69	5.79	20.12	33.65
	SH*	32.28	34.92	10.59	49.17	4.51	11.95	33.67

*LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

Table 3: The % difference between implementing factor-levels of the MC methodology, for different environmental factors, for each KPI (a) – (g)

When comparing the benchmark policy of no menus created (b = 1) to the MC methodology with either a smaller menu size of b = 3 or a larger menu size of b = 5, we observe a positive, and statistically significant increase in the % change values for all KPIs (i) – (vii). Thus, implementing a menu creation method has the potential to significantly improve both NPO and volunteer KPIs. A larger menu size of 5 provides more task selection options for volunteers, which results in more options to create better final group creations and assignments. Thus, when the benchmark policy of no menu created (b = 1) is compared to MC methodology with a larger menu size of b = 5, a positive, and statistically significant, increase is observed in all KPIs (i) – (vii). The results comparison of implementing a menu of size of b = 3 vs. implementing a menu of size b = 5 is presented in Appendix X9. Yet, it is observed that volunteer behavior does influence the magnitude of these results. When a NPO has High Picky Volunteers (HPV), providing them with a larger menu of tasks (b = 5) to pick from leads to the most improvement being achieved compared to traditional methods of assigning a single task to each volunteer (b = 1), with an average increase of 54.86% in KPI (i) - 'objective function value'.

The KPIs show an improvement for all environmental settings for the smaller menu size (b = 3), with some exception. The exception is KPI (v) - '% of volunteer pairs with similar age IDs', where a decrease in KPI performance is observed for the environmental settings High Picky Volunteers (HPV) (-2.37%), and volunteers with No Heterogeneity (NH) (-0.50%). This demonstrates that environmental instances where the NPO is dealing with volunteers that are pickier about the tasks they select from menus, and instances where volunteer's task selection behaviors have no variability, a smaller menu of size 3 provides lesser options for the volunteers to select from, resulting in lesser task selections and ultimately, a decrease in the % of groups homogenous by age ID created. A larger menu size of 5 provides more task selection options for volunteers, which results in final group creations and assignments where the age ID similarity among volunteers within created groups is better met. Overall, to achieve better group creations and assignments in environments with High Picky Volunteers and No Heterogeneity volunteers, it is beneficial to have larger task recommendation menus.

6.2 Impact of Group Creation Constraints

In this section, we present the average percent change in KPI values when among different factor levels the Group Creation methodology is implemented. In the GC methodology, the benchmark policy is when skill difference d = 7 and age ID difference a = 7 are implemented, i.e., when the group creation (GC) constraints for the creation of ideal and homogenous groups are not implemented (maximum possible value of difference in skill levels and age IDs among volunteers is 7). In Table 4, we present the % change between Not implementing ideal age ID difference among group members (a) and ideal skill difference among group members (d) in our GC constraints (a = 7, d = 7) vs. Implementing these GC constraints (a = 1, d = 1). We also present the % change between Not implementing GC constraints (a = 7, d = 7) vs. when only one or the other GC constraint is implemented (either a = 1, d = 7 or a = 7, d = 1).

We do not present the KPI (i) - Objective function value as it captures penalties based on the specific d and a values used, and thus, does not provide for an accurate comparison measure of evaluating the impact of implementing or not implementing group creation constraints. Instead, we examine KPIs '% of volunteer pairs with similar age IDs' (v) and '% of volunteers assigned in groups with similar skill level values' (vi) in detail to understand impact of implementing ideal group creation constraints.

Avg. Change (%)	**Environmental Factors**	**KPIs**					
		(ii)	**(iii)**	**(iv)**	**(v)**	**(vi)**	**(vii)**
Ideal Group Creation Cts Not Implemented (7,7) vs. Implemented (1,1)							

d = 7 a = 7 vs. d = 1 a = 1	**Avg Over All Env. Instances**	-12.04	0.68	-16.15	36.83	15.30	-3.63
	p-value	<<0.05	0.17	<<0.05	<<0.05	<<0.05	<<0.05
	LPV*	-12.30	-0.29	-14.92	45.59	21.71	-4.70
	HPV*	-10.39	2.13	-17.12	28.74	6.22	-1.93
	NH*	-13.44	0.93	-19.01	30.03	10.21	-4.41
	SH*	-11.72	0.51	-15.08	44.07	19.18	-3.10
Ideal Group Creation Cts Not Implemented (7,7) vs. Implement Age ID difference a (7,1)							
d = 7 a = 7 vs. d = 7 a = 1	**Avg Over All Env. Instances**	-3.37	-0.17	-4.18	34.24	-2.44	-0.41
	p-value	<<0.05	0.57	<<0.05	<<0.05	<<0.05	0.31
	LPV*	-3.77	-0.21	-4.46	46.92	-1.93	-0.30
	HPV*	-2.65	-0.11	-3.62	22.72	-3.17	-0.57
	NH*	-4.15	-0.52	-5.70	24.77	-4.42	-0.81
	SH*	-2.88	0.07	-3.28	44.31	-0.94	-0.13
Ideal Group Creation Cts Not Implemented (7,7) vs. Implement Skill level difference d (1,7)							
d = 7 a = 7 vs. d = 1 a = 7	**Avg Over All Env. Instances**	-7.71	0.30	-10.43	3.00	17.41	-4.26
	p-value	<<0.05	0.57	<<0.05	0.45	<<0.05	0.11
	LPV*	-7.32	-0.79	-8.77	0.61	23.51	-4.69
	HPV*	-5.87	1.91	-10.54	4.17	11.54	-1.27
	NH*	-11.86	0.10	-16.22	4.44	10.86	-9.41
	SH*	-6.38	0.44	-8.45	1.61	20.84	-2.31

*LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

Table 4: The % difference between implementing and not implementing ideal age ID difference and skill difference constraints, for different environmental factors, for each KPI (ii) – (vii)

In Table 4, we first look at the average percent difference (over all environmental instances) of implementing the group creation constraints to create ideal volunteer groups (d = 1, a = 1) versus not implementing the constraints (d = 7, a = 7). As expected, there is a statistically significant increase in KPIs capturing the impact on group creation, i.e., increase of 36.83% in KPI (v) - '% of volunteer pairs with similar age IDs', and 15.30% in KPI (vi) - '% of volunteers assigned in groups with similar skill level values', when group creation constraints are implemented. Yet, there is also a statistically significant decrease of 12.04% in KPI (ii) – 'total number of volunteers assigned', a decrease of 16.15% in KPI (iv) – 'number of volunteers assigned in groups', and a decrease of 3.63% in KPI (vii) – '% of total task-skill requirements met' is observed. This illustrates that a NPO would face a tradeoff of loss in performance in volunteer and NPO KPIs (ii), (iv), (vii) when trying to increase the number of ideal group creations and task assignments being made. This tradeoff was observed to be consistent across the different and environmental instances of our DOE (*LPV, HPV, NH, SH).

Next, we look at insights when only one of the ideal group creation constraints is implemented (either ideal age ID difference or ideal skill difference, but not both being implemented together). When only the age id difference constraint with a = 1 is implemented, as expected, this significantly increases KPI (v) - '% of volunteer pairs that have a similar age ID' by 34.24%, but a statistically significant decrease is observed for KPIs (ii) – Total number of volunteers assigned, (iv) – 'number of volunteers assigned in groups', and (vi) – '% of volunteers assigned in groups with similar skill level values'.

We also look at the percent change in KPIs when only the skill level difference constraint d = 1 is implemented. A decrease in most KPI values is observed, but as expected, a significant increase of 17.41% in KPI (vi) - '% of volunteers assigned in groups with similar skill level values' is observed. There is a non-significant increase of 3.00% in KPI (v) - '% of volunteer pairs with similar age IDs'. This concludes that implementing the skill level difference constraint aids in creating more groups where age IDs are also similar, but this is not a significant increase. This observation is supported by the findings in volunteer behavior in our empirical data analysis in Figure 4 (Section 4). The age ID of a volunteer influences which task a volunteer is willing to work on, with younger volunteers (age ID 1-5) wanting to work on a wide variety of more complex online tasks, while the older volunteers (age ID 6-7) select only a few types of simple online tasks (e.g., data entry, subtitle writing). As the majority of the volunteers completing online tasks are young and within age IDs 1-5 (Section 4, Figure 3), and display willingness to work on a larger variety of tasks at the NPO, implementing the constraint for obtaining skill similarity within groups also results in achieving groups that have volunteers with similar age IDs in them.

Overall, when comparing the implementation of only the skill level difference constraint d= 1 versus only implementing the age id difference constraint a = 1, we observe that KPIs (ii) – 'Total number of volunteers assigned', (iv) – 'Number of volunteer assigned in groups' perform better. A NPO with similar volunteer characteristics as our empirical data collection is better off implementing the age ID similarity among volunteers rather than skill level similarity among volunteers for group creation.

Next, in Table 5, we present the average percent change in the values when minimum group size is not implemented ($\hat{g}_r$ = 1) (when the assignment of a single volunteer to a task is not penalized) vs. a minimum group size of 2 is implemented ($\hat{g}_r$ = 2).

Avg. Change (%)	**Environmental Factors**	**KPIs**					
		(ii)	**(iii)**	**(iv)**	**(v)**	**(vi)**	**(vii)**
Min Group Size Not Implemented (1), Implemented (2)							
$\hat{g}_r$ = 1 vs. $\hat{g}_r$ = 2	**Avg Over All Env. Instances**	0.03	-0.79	1.67	-2.17	-0.06	-1.10
	p-value	0.57	<<0.05	0.61	<<0.05	0.43	0.21
	LPV*	1.09	0.07	2.88	-0.91	1.91	1.56
	HPV*	-1.13	-2.03	0.18	-3.31	-1.86	-3.91
	NH*	-0.05	-0.77	2.18	-5.12	-1.12	-2.12
	SH*	-0.24	-0.79	1.02	0.96	0.11	-1.13

*LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

Table 5: The % difference observed upon implementing and not implementing minimum group size constraint, for different environmental factors, for each KPI (ii) – (vii)

When a minimum group size constraint is applied ($\hat{g}_r = 2$), a penalty is imposed for assigning single volunteers to tasks. In Table 5, we compare the percent difference in KPI values when a minimum group size is implemented vs. when it is not. A negative % difference value indicates a worse performance of the KPI when a minimum group size is implemented. Only the KPIs (iii) – 'Total number of tasks with volunteer(s) assigned' (0.79%) and (v) – '% of volunteer pairs with similar age IDs' (2.17%) display statistically significant decreases. Other KPIs do not show statistically significant changes, indicating that implementing a penalty for failing to meet the minimum group size in task assignments does not heavily impact most performance metrics. This suggests that penalizing assignments that do not make groups of two or more could be beneficial for NPOs, as it encourages group-based task assignments desired by volunteers, without significantly affecting most key performance indicators.

6.3 Model Computational Time

All optimization models in our methodology have been implemented in IBM ILOG CPLEX Optimization Studio IDE 20.1.0. We run experiments on computers with an Intel Core i7-8650U CPU @ 1.9 GHz processor, RAM of 32 GB, and a 64-bit OS, x64-based processor. Solve time to solve whole methodology, all in CPLEX. For all scenarios in MC and to obtain final GC result. Table 6 presents the runtime statistics (mean, minimum, maximum, standard deviation (SD)) in minutes for each experiment run (1 replication). The solve time presented captures the time for the entire methodology i.e., to run all scenarios runs for the MC IP, plus to run the Consensus IP, and the GC IP. The results are presented across all experiments, and also conditioned on each factor level value. The mean runtime increases significantly with increase in menu size, as the platform has more options to consider when creating volunteer task recommendation menus, and more options to consider when creating volunteer groups and task assignments after collecting multiple volunteer task selections. The mean run time is the highest when the ideal group creation constraint for obtaining volunteer groups with similar skill level values is implemented (when d = 1) as the model must implement a restricting constraint to identify group creations where skill levels of volunteers assigned in a group are similar. Furthermore, mean runtime is highest when volunteer pickiness is low (*LPV) and when volunteer displays some heterogeneity (*SH) in task selection behaviors because there are more tasks with willing volunteers, increasing the options that the model has for finding ideal volunteer for group creation and assignment.

Factor	Level	Mean	Min	Max	SD
Average Over All Factors		27.77	2.61	273.39	24.68
Min Grp Size $\hat{g}_r$	1	27.23	2.61	259.82	22.05
	2	28.44	4.33	273.39	20.94
Max Menu Size b	1	15.98	2.61	71.59	8.58
	3	21.60	2.76	182.45	14.88
	5	49.04	3.24	273.39	32.40
Maximum desired skill difference among volunteers in groups (d)	7	7.60	2.61	23.74	2.21
	1	48.94	4.33	273.39	28.72
Maximum desired age ID difference among volunteers in a group (a)	7	30.46	4.17	273.39	22.89
	1	24.81	4.17	268.03	19.94
α_{0v} generation	*LPV, NH	48.09	2.61	268.03	31.30
	*HPV, NH	5.14	4.17	6.34	0.30
	*LPV, SH	41.66	7.59	273.39	23.86
	*HPV, SH	5.12	4.17	5.92	1.08

*LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

Table 6: Runtime Statistics (in Minutes) for 1 Replication run with 24 Tasks and 72 Volunteers

7. Conclusions, Discussions and Future Research

This work was motivated by an empirical data analysis of a 2023 volunteer on-boarding dataset. This NPO uses remote volunteers to work on online content creation tasks. This analysis highlighted that demographic information can serve as valuable data for making volunteer-to-task assignments. For instance, our analysis revealed that younger volunteers often selected a broader variety of tasks, driven by a desire to learn and gain experience, while older volunteers tended to prefer simpler tasks, motivated by a commitment to contribute to the organization while also seeking opportunities to network and engage with the volunteer community. This observation is aligned with findings in existing literature (Bang, 2015; Harrison, 1995). Findings from our empirical data analysis can be used to inform further research on methodologies developed to make volunteer management decisions. Probability distributions derived from our partner NPO's volunteer dataset can assist in generating representative data for volunteers participating in nonprofit organizations, reflecting their specific behaviors and preferences on crowdsourcing platforms.

This empirical analysis motivated our developed methodology, which creates personalized task menus from which volunteers have autonomy to select tasks they are willing to on, and then creates groups of willing volunteers with ideal group characteristics (i.e., their collective skills meet NPO task's needs, and there is learning potential and affinity among volunteers in groups to encourage volunteer learning, networking, and engagement). The methodology consists of solving three integrated integer programs, one for menu creation (MC Method), one to determine the consensus of the menus generated in multiple scenarios (Consensus IP), and one for volunteer group creation and task assignments (GC Method). First, the MC method creates personalized task recommendation menus offered to volunteers that requires a volunteer's willingness to work on task(s) to be estimated. Yet, because individual volunteer preferences are stochastic, we solve this MC method for multiple scenarios of volunteer willingness data. Next, a consensus integer program is implemented to synthesize the task offerings on menus obtained across the multiple scenarios into a final task recommendation menu that is offered to each volunteer. Volunteer's willingness to work on task(s) offered on their menu is then captured, which is then used as an input into the GC model to create volunteer groups and make task assignments. Volunteer groups are created based on the learning potential and affinity levels among volunteers in groups to encourage volunteer learning, networking and engagement while also aiming to assign groups that have the necessary skill levels to complete tasks. The overall objective of the methodology is to maximize the number of volunteers assigned to tasks while incorporating penalties if volunteer to task assignments do not meet minimum/maximum group size, task's skill level requirement(s), desired skill difference among volunteers within a group, desired age ID difference among volunteers in a group, and size of menu of tasks offered to volunteer(s). Implementing the menu creation methodology increased the objective function values on average by 34.62%, providing a statistically significant improvement to nonprofits over the traditional method of offering a single task to each volunteer. This approach increased the total number of volunteers assigned, the number of tasks with volunteers assigned, the number of volunteers assigned in groups, and the number of ideal groups formed. A menu creation method is especially useful when a NPO has highly picky volunteers, with average objective function values increasing by 54.86% over the benchmark of offering a single task to each volunteer. Furthermore, an NPO would face a tradeoff when trying to increase the number of ideal group creations. While KPIs associated with ideal group creation increased, the number of total volunteers assigned and the number of tasks with volunteers decreased.

Our static optimization modeling framework can be run iteratively by NPOs to create and assign volunteer groups to tasks. However, it does not account for the impact and interrelatedness of current volunteer assignments on future volunteer retention. Urrea & Yoo, 2021 highlight that well-matched tasks reduce volunteer attrition and increase future retention (Urrea & Yoo, 2021). While dynamic models for worker-to-task assignments in for-profit contexts exist (Ausseil et al., 2022; Horner et al., 2021; Horner et al., 2024), further research is needed to develop dynamic models to capture the impact of initial volunteer to group assignments on future volunteer satisfaction, engagement, and long-term retention. Qualitative literature suggests that creation of volunteer groups to work on NPO tasks is a beneficial approach to achieve volunteer satisfaction, retention and NPO goals (Jensen & McKeage, 2015; Millette & Gagné, 2008). Thus, the development of a dynamic model for creation and assignment of volunteer groups to NPO tasks would be a meaningful next step. This can be done by creating and integrating prediction models that utilize the impact of implementing volunteer management theories on volunteer and NPO motivations and behaviors. Qualitative studies highlight that volunteers in both traditional NPOs and volunteer crowdsourcing platforms have diverse behaviors, interests, commitment levels and performance capabilities, which influences successful NPO task completion (Macduff, 2005; Urrea & Yoo, 2021). Models can account for different volunteer arrival behaviors, such as periodic, episodic, virtual, or spontaneous, and classify volunteers by experience level (beginner, intermediate, advanced) and dedication (dedicated, non-dedicated) (Urrea & Yoo, 2021). NPOs also have various tasks (e.g., long-term, continuous, episodic, urgent, non-urgent) (Macduff, 2005; Urrea & Yoo, 2021). Thus, identifying and assigning the right type of volunteer to the right task influences productivity, experience, retention, and task completion in NPOs (Urrea & Yoo, 2021). Our model, designed to create homogeneous volunteer groups with high affinity and peer-learning potential, could be expanded to consider these volunteer and task characteristics to obtain better volunteer group to task assignments in NPOs.

Another limitation of static models is their inability to monitor the progress of complex tasks that require multiple periods to be completed successfully (e.g., video editing), thus another valuable research direction is the development of models designed to monitor and support the completion of complex NPO tasks. Current literature reviewed for NPO applications focuses on dynamic optimization models for simple, one-time tasks (e.g., pick-up and drop-off food delivery from NPO) (Manshadi & Rodilitz, 2021; Schmidt & Albert, 2022; Shi et al., 2021), or for task assignment of self-selected volunteer groups to tasks (group not created from individual volunteers arriving) (Chen, Wang, et al., 2021; Chen, Zhang, et al., 2021). Thus, future work can develop dynamic models that create volunteer groups and assigns them to long-term, complex tasks, where both volunteers and task arrivals change overtime. One approach could be to capture task status and volunteer participation history at the start of each period to identify ideal assignments based on task needs and volunteer reputation, determined by their past participation at the NPO. Reviewing models developed for worker reputation tracking in for-profit organizations can provide insights for implementing a similar system for NPOs.

While our methodology was designed for NPOs that consistently provide services, it could be adapted for NPOs involved in disaster response. This would require further developments to address the unique challenges and uncertainties faced in providing emergency support services, compared to delivering routine, ongoing support to communities. Traditional NPOs and disaster response NPOs differ significantly in both their organizational setup and how they mobilize their resources. Furthermore, both types of NPOs attract volunteers with different profiles and motivations, resulting in distinct volunteer management needs. Long-term volunteers often require structured and ongoing engagement, enabling sustained support and deeper integration into the organization's culture (Tsai et al., 2023). In contrast, disaster response organizations must quickly mobilize spontaneous volunteers, adapting to high uncertainty and rapid changes in demand (Tsai et al., 2023). Spontaneous volunteers—who join efforts without formal prior affiliation—play a vital role during emergencies but often require more coordination and structure as the situation evolves (Carius et al., 2024). Thus, an interesting future research direction is to expand this work to disaster response scenarios with dynamic resource allocation and real-time volunteer management.

Acknowledgements

This research was partially funded by the US National Science Foundation CAREER award #1751801

References

Alam, S. L., & Campbell, J. (2017). Temporal Motivations of Volunteers to Participate in Cultural Crowdsourcing Work. *Information Systems Research*, *28*(4), 744–759. https://doi.org/10.1287/isre.2017.0719

Ausseil, R., Pazour, J. A., & Ulmer, M. W. (2022). Supplier Menus for Dynamic Matching in Peer-to-Peer Transportation Platforms. *Transportation Science*, *56*(5), 1304–1326. https://doi.org/10.1287/trsc.2022.1133

Bang, H. (2015). Volunteer age, job satisfaction, and intention to stay: A case of nonprofit sport organizations. *Leadership & Organization Development Journal*, *36(2)*, 161–176.

Barnes, M. L., & Sharpe, E. K. (2009). Looking Beyond Traditional Volunteer Management: A Case Study of an Alternative Approach to Volunteer Engagement in Parks and Recreation. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *20*(2), 169–187. https://doi.org/10.1007/s11266-009-9080-5

Barron, P., & Rihova, I. (2011). Motivation to volunteer: A case study of the Edinburgh International Magic Festival. *International Journal of Event and Festival Management*, *2*(3), 202–217. https://doi.org/10.1108/17582951111170281

Baruch, A., May, A., & Yu, D. (2016). The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform. *Computers in Human Behavior*, *64*, 923–931. https://doi.org/10.1016/j.chb.2016.07.039

Berenguer, G., & Shen, Z.-J. (Max). (2020). Challenges and Strategies in Managing Nonprofit Operations: An Operations Management Perspective. *Manufacturing & Service Operations Management*, *22*(5), 888–905. https://doi.org/10.1287/msom.2018.0758

Carius, C., Graw, J., & Schultz, C. (2024). *Achieving Organizationality in Large-Scale Crises: A Comparative Case Study on the Communicative Constitution of Spontaneous Volunteer Collectives*. https://doi.org/10.1177/08997640241233726

Chen, S.-Q., Wang, Y.-M., Shi, H.-L., & Zhang, X.-X. (2021). A decision-making method for uncertain matching between volunteer teams and rescue tasks. *International Journal of Disaster Risk Reduction*, *58*, 102138. https://doi.org/10.1016/j.ijdrr.2021.102138

Chen, S.-Q., Zhang, L., Shi, H.-L., & Wang, Y.-M. (2021). Two-sided matching model for assigning volunteer teams to relief tasks in the absence of sufficient information. *Knowledge-Based Systems*, *232*, 107495. https://doi.org/10.1016/j.knosys.2021.107495

Deloitte Volunteer Impact Research. (2017). https://www2.deloitte.com/us/en/pages/about-deloitte/articles/citizenship-deloitte-volunteer-impact-research.html

Escallon-Barrios, M., Noham, R., & Smilowitz, K. (2024). Dual mode scheduling in volunteer management. *Socio-Economic Planning Sciences*, *92*, 101796.

Esfandiari, M., Wei, D., Amer-Yahia, S., & Basu Roy, S. (2019). Optimizing peer learning in online groups with affinities. *International Conference on Knowledge Discovery & Data Mining*, 1216–1226.

Fait, M., & Sakka, G. (2020). Knowledge sharing: An innovative organizational approach to engage volunteers. *EuroMed Journal of Business*, *16(3)*, 290–305.

Falasca, M., & Zobel, C. (2012). An optimization model for volunteer assignments in humanitarian organizations. *Socio-Economic Planning Sciences*, *46*(4), 250–260. https://doi.org/10.1016/j.seps.2012.07.003

Fullwood, R., & Rowley, J. (2021). The role of knowledge sharing in volunteer learning and development. *Nonprofit Management and Leadership*, *32*(1), 121–139. https://doi.org/10.1002/nml.21466

Garcia, C., Rabadi, G., & Handy, F. (2018). Dynamic resource allocation and coordination for high-load crisis volunteer management. *Journal of Humanitarian Logistics and Supply Chain Management*. https://doi.org/10.1108/JHLSCM-02-2018-0019

Berenguer, G., Haskell, WB, & Li, L (2023). Managing Volunteers and Paid Workers in a Nonprofit Operation. *Management Science*, *70*(8), 5298–5316.

Harrison, D. A. (1995). Volunteer motivation and attendance decisions: Competitive theory testing in multiple samples from a homeless shelter. *Journal of Applied Psychology*, *80(3)*.

Henderson, S. G., Berg, P. L. V. D., Jagtenberg, C. J., & Li, H. (2022). How should volunteers be dispatched to out-of-hospital cardiac arrest cases? *Queueing Systems*, *100(3-4)*, 437–439.

Horner, H., Pazour, J. A., & Mitchell, J. E. (2024). Increasing Driver Flexibility through Personalized Menus and Incentives in Ridesharing and Crowdsourced Delivery Platforms. *Naval Research Logistics*.

Horner, H., Pazour, J., & Mitchell, J. E. (2021). Optimizing driver menus under stochastic selection behavior for ridesharing and crowdsourced delivery. *Transportation Research Part E: Logistics and Transportation Review*, *153*, 102419. https://doi.org/10.1016/j.tre.2021.102419

Jensen, K. B., & McKeage, K. K. (2015). Fostering Volunteer Satisfaction: Enhancing Collaboration Through Structure. *The Journal of Nonprofit Education and Leadership*, *5*(3), n/a.

Kaur, M. P., Smith, S., Pazour, J. A., & Duque Schumacher, A. (2022). Optimization of volunteer task assignments to improve volunteer retention and nonprofit organizational performance. *Socio-Economic Planning Sciences*, *84*, 101392. https://doi.org/10.1016/j.seps.2022.101392

Kazemi, L. (2012). GeoCrowd: Enabling query answering with spatial crowdsourcing. *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, 189–198.

Lo, I., Manshadi, V., Rodilitz, S., & Shameli, A. (2021). Commitment on Volunteer Crowdsourcing Platforms: Implications for Growth and Engagement. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3802628

Macduff, N. (2005). Societal changes and the rise of the episodic volunteer. *Emerging Areas of Volunteering*, *1*(2), 49–61.

Manshadi, V., & Rodilitz, S. (2021). *Online Policies for Efficient Volunteer Crowdsourcing* (SSRN Scholarly Paper ID 3802624). Social Science Research Network. https://doi.org/10.2139/ssrn.3802624

Mazlan, N., Syed Ahmad, S. S., & Kamalrudin, M. (2018). Volunteer selection based on crowdsourcing approach. *Journal of Ambient Intelligence and Humanized Computing*, *9*(3), 743–753. https://doi.org/10.1007/s12652-017-0490-8

Millette, V., & Gagné, M. (2008). Designing Volunteers' Tasks to Maximize Motivation, Satisfaction and Performance: The Impact of Job Characteristics on Volunteer Engagement. *Motivation and Emotion*, *32*, 11–22. https://doi.org/10.1007/s11031-007-9079-4

Minn, S. Y., Y., D., M. C., Z., & Vie, J. J. (2018). Deep knowledge tracing and dynamic student classification for knowledge tracing. *2018 IEEE International Conference on Data Mining (ICDM)*, 1182–1187.

Mofidi, S. S., & Pazour, J. A. (2019). When is it beneficial to provide freelance suppliers with choice? A hierarchical approach for peer-to-peer logistics platforms. *Transportation Research Part B: Methodological*, *126*, 1–23. https://doi.org/10.1016/j.trb.2019.05.008

NCCS Team. (2022, March 22). The Nonprofit Sector in Brief—2019. *National Center for Charitable Statistics*. https://nccs.urban.org/publication/nonprofit-sector-brief-2019

Rahman, H., Roy, S. B., Thirumuruganathan, S., Amer-Yahia, S., & Das, G. (2019). Optimized group formation for solving collaborative tasks. *The VLDB Journal*, *28*(1), 1–23. https://doi.org/10.1007/s00778-018-0516-7

Rauchecker, G., & Schryen, G. (2018, May). *Decision Support for the Optimal Coordination of Spontaneous Volunteers in Disaster Relief*. 15th International Conference on Information Systems for Crisis Response and Management, Rochester, NY, USA. https://epub.uni-regensburg.de/37001/

Sánchez, O. R., Ordonez, C. A., Duque, M. A. R., & Pinto, I. I. B. S. (2021). Homogeneous group formation in collaborative learning scenarios: An approach based on personality traits and genetic algorithms. *IEEE Transactions on Learning Technologies*, *14*(4), 486–499.

Schmidt, A. P., & Albert, L. A. (2022). Task recommendations for self-assigning spontaneous volunteers. *Computers & Industrial Engineering*, *163*, 107798. https://doi.org/10.1016/j.cie.2021.107798

Shi, Z. R., Lizarondo, L., & Fang, F. (2021a). A Recommender System for Crowdsourcing Food Rescue Platforms. *Proceedings of the Web Conference 2021*, 857–865. https://doi.org/10.1145/3442381.3449787

Shi, Z. R., Lizarondo, L., & Fang, F. (2021b). A Recommender System for Crowdsourcing Food Rescue Platforms. *Proceedings of the Web Conference 2021*, 857–865. https://doi.org/10.1145/3442381.3449787

Tongarlak, A. B., Lee, M. H., & Field, J. (2024). A dynamic model for managing volunteer engagement. *Operations Research*.

Tsai, A. C. Y., Newstead, T., Lewis, G., & Chuah, S.-H. (2023). *Leading Volunteer Motivation: How Leader Behaviour can Trigger and Fulfil Volunteers' Motivations*. *35*, 266–276.

Urrea, G., & Yoo, E. (2021). The Role of Volunteer Experience on Performance on Online Volunteering Platforms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3784152

Vinella, F. L., Hu, J., Lykourentzou, I., & Masthoff, J. (2022). Crowdsourcing Team Formation With Worker-Centered Modeling. *Frontiers in Artificial Intelligence*, *5*, 818562. https://doi.org/10.3389/frai.2022.818562

Wei, D., Koutis, I., & Roy, S. B. (2021). *Peer learning through targeted dynamic groups formation*. 121–132.

Wiens, P. D., Zizzi, C., & Heatwole, C. (2022). Instructional Grouping Theory: Optimizing Classrooms and the Placement of Ranked Students. *Educational Practice and Theory*, *5–21*(44(1)).

Zhen, Y., Khan, A., & Nazir, S. (2021). Crowdsourcing usage, task assignment methods and crowdsourcing platforms: A systematic literature review. *Journal of Software: Evolution and Process*, *33(8)*.

Zheng, Y., Li, C., Liu, S., & Lu, W. (2018). An improved genetic approach for composing optimal collaborative learning groups. *Knowledge-Based Systems*, *139*, 214–225.

Appendix

X1 - E_{rs}(skill level required for a skill s, on task r) values defined for design of experiments using understanding of NPO projects, tasks and their needs. We then map these skills for each of the 12 online tasks (e.g., a web development task would require skills S1-coding and development, and S2-basic computer skills).

Task ID	Task	Skill ID	Skill Name	Skill Level Needed
1	Data Entry	S2	Basic Computer Skills	5
2	Blog Writing	S2	Basic Computer Skills	2
		S3	Languages	2
		S4	Writing	2
3	Pamphlet Creation	S2	Basic Computer Skills	3
		S3	Languages	3
		S4	Writing	2
4	Animation	S2	Basic Computer Skills	3
		S5	Animation	3
5	Web Development	S1	Coding & Development	5
		S2	Basic Computer Skills	2
6	App Development	S1	Coding & Development	5
		S2	Basic Computer Skills	2
7	Video Editing	S2	Basic Computer Skills	3
		S6	Video Editing	2
8	Translation	S2	Basic Computer Skills	1
		S3	Languages	4
		S4	Writing	2
		S7	Language Fluency	4
9	Subtitle Writing	S2	Basic Computer Skills	1
		S3	Languages	4
		S4	Writing	3
		S7	Language Fluency	4
10	Graphic Design	S2	Basic Computer Skills	3
		S8	Graphic Design	2
11	Social Media Manager	S2	Basic Computer Skills	3
		S3	Languages	2
		S4	Writing	1
		S9	Social Media Management	2

		S10	Analytics	3
12	Audio Content Manager	S2	Basic Computer Skills	3
		S9	Social Media Management	2

Table: Task skill level requirement values (E_{rs}) values defined by observing NPO task requirements

X2 - C_{vs}(skill level values of volunteer(s) in skill(s), extracted from NPO's volunteer dataset based on the following terminologies extracted from volunteer's open-ended responses to questions in preliminary survey collected.

Skill	Terms extracted to determine distribution of volunteer's skill levels in each skill
S1 – Coding and development	Software, computer, IT, engineering, code, web design, development, technology
S2 – Basic computer skills	Computer, data, school, student, college, university, engineering, technology
S3 – Languages	Sum of languages fluent in, English, Punjabi, Hindi, other language(s)
S4 – Writing	Sum of languages fluent in, Write (other languages familiar with), Fluent, Advanced, Write (Career and Education Qualifications), Write(Past Volunteering Experience),
S5 – Animation	Computer, Animation, Design, college, university
S6 – Video Editing	Computer, Design, Vlogging, Social Media, college, university
S7 – Language Fluency	Sum of languages fluent in, Write, Read, Fluent, Advanced, Speak
S8 – Graphic Design	Web design, design, college, university
S9 – Social Media Management	Manage, Media, Social Media, Manage, Vlogging
S10 – Analytics	Data, Analytics, Business, Manage

Table: Volunteer skill level values (C_{vs}) defined by extracting particular terminologies from volunteer's open-ended responses to onboarding survey questions

X3 - Upon observing the mean and standard deviation (SD) of the objective function values of experiments run with 5 replications and 10 replications, it is noted that there is not much difference in the mean of objective function values obtained for 5 replications versus 10 replications (0.3) and the SD (0.5). Thus, to reduce computational time while still obtaining similar results, we choose to run experiments with 5 replications.

10 Scenarios	**5 Replications**	**10 Replications**
Mean	5.2	4.9
SD	7.1	6.6

Table: Mean and SD of objective function values when 5 replications vs. 10 replication are run per experiment

Next, we observe objective function results obtained for running 5 replications with 5 scenarios vs. 5 replications with 10 scenarios in the menu creation methodology. It is observed that there is not much difference between the means of the two experiments (1.0), but there is a high SD observed for the objective function values (11.6) for experiments with 5 scenarios, while a lower SD (7.1) was observed for the experiment with 10 scenarios. Thus, we choose to run each experiment with 5 replications and 10 scenarios to ensure the consideration of ideal number of situations to model the volunteer's variable task selection behaviors.

5 Replications	**5 Scenarios**	**10 Scenarios**

Mean	**-15**	**-16**
SD	**11.6**	**7.1**

Table: Mean and SD of objective function values when 5 scenarios vs. 10 scenarios are run per experiment

X4 - Formula for rescaling Objective function values KPI

X = Current objective function value

$from\_max$ = Maximum value of current objective function values

$from\_min$ = Minimum value of current objective function values

to_max = Maximum value of new desired range of values = 100

to_min = Minimum value of new desired range of values = 0

$$New\ objective\ function\ value = \frac{(X - from\_min) * (to\_max - to\_min)}{(from\_\max - from\_\min) + to\_min}$$

X5 - The conditional probability distributions of task selections for different volunteer age ID categories

Age ID	Age Range	Task ID	Task Name	Distribution of task being selected
1	<15	3	Pamphlet Creation	0.40
		5	Web Development	0.17
		6	App Development	0.20
		7	Video Editing	0.23
2	15-19	1	Data Entry	0.19
		2	Blog Writing	0.27
		3	Pamphlet Creation	0.06
		4	Animation	0.03
		5	Web Dev	0.20
		6	App Dev	0.09
		7	Video Editing	0.05
		8	Translation	0.05
		9	Subtitle Writing	0.04
		12	Audio Content Mgmt	0.03
3	20-24	1	Data Entry	0.19
		2	Blog Writing	0.19
		4	Animation	0.04
		5	Web Dev	0.16
		6	App Dev	0.07
		7	Video Editing	0.15
		8	Translation	0.13
		9	Subtitle Writing	0.04
		12	Audio Content Mgmt	0.05

4	25-29	1	Data Entry	0.26
		2	Blog Writing	0.22
		3	Pamphlet Creation	0.08
		5	Web Dev	0.14
		7	Video Editing	0.10
		8	Translation	0.08
		9	Subtitle Writing	0.13
5	30-34	1	Data Entry	0.30
		2	Blog Writing	0.30
		5	Web Dev	0.10
		7	Video Editing	0.30
6	35-39	1	Data Entry	0.33
		8	Translation	0.33
		9	Subtitle Writing	0.33
7	40-44	1	Data Entry	0.30
		2	Blog Writing	0.40
		3	Pamphlet Creation	0.10
		7	Video Editing	0.20

X6 - Probability distribution of volunteers of different skill levels in skills S1 – S10

Skill Number	Skill	Skill Levels						
		1	2	3	4	5	6	7
S1	Coding & Dev	0.655	0.177	0.108	0.06	0	0	0
S2	Basic Computer Skills	0.437	0.232	0.152	0.08	0.075	0.024	0
S3	Languages	0	0.099	0.432	0.283	0.136	0.05	0
S4	Writing	0.028	0.456	0.214	0.147	0.091	0.064	0
S5	Animation	0.721	0.169	0.11	0	0	0	0
S6	Video Editing	0.629	0.179	0.092	0.074	0.026	0	0
S7	Language Fluency	0.014	0.465	0.211	0.155	0.07	0.07	0.015
S8	Graphic Design	0.661	0.146	0.108	0.065	0.02	0	0
S9	Social Media Mgmt	0.529	0.099	0.128	0.094	0.115	0.035	0
S10	Analytics	0.659	0.085	0.146	0.11	0	0	0

The probability distribution of volunteers with different levels of skills

X7 – ANOVA Tests to identify interactions between factors ad levels implemented on the KPIs (response variables) being collected

In this appendix section we present ANOVA results in Tables A5 – A11, to observe the interaction between the multiple factors implemented in our design of experiments (minimum group size, maximum menu size, volunteer pickiness, volunteer heterogeneity, skill similarity implemented, age ID similarity implemented) with the 7 different response variables (KPIs) that we collect as a part of our experiments. We present all the main factors, as well as the second order interactions. For ANOVA, we use a significance level of 0.05, and the factors denoted by (*) have a p-value less than 0.05. To meet the assumptions of ANOVA, we first normalize the response variable being considered.

We see that there is a significant main effect impact on each response variable (KPI (a) – (g)), for almost all factors implemented in our design of experiments. The KPIs Objective function value (a), Number of volunteers assigned in groups (d) and, % of volunteers assigned in groups with similar skill level values (f), have a significant main effect impact from almost all factors implemented. Thus, as expected nonprofits can achieve higher objective function values with less picky volunteers than ones who are highly picky. With less picky volunteers, they also have the ability to increase the number of volunteers assigned in groups and create more ideal groups. Similarly, a NPO benefits from variation in volunteer preferences for tasks, which results in higher objective function values, more groups, and more ideal groups.

We further note that the factor minimum group size does not have a significant main effect impact on the KPIs total number of volunteers assigned (b), total number of tasks with volunteer assignments (c), % of total task-skill requirements met (g), which indicates that restricting assignments to abide by a minimum group size constraint does not significantly impact these KPIs important to a NPO for successful task completion. The factor desired skill level difference among volunteers has a main effect impact on KPIs has a main effect impact on KPIs Objective function value (a), Total number of volunteers assigned (b), Number of volunteers assigned in groups (d), % of volunteers assigned in groups with similar skill level values (f), % of total task-skill requirements met (g), but not on KPIs does not have a main effect impact on KPIs total number of tasks with volunteer assignments (c), % of volunteer pairs with similar age IDs (e).The factor desired age ID difference among volunteers has a main effect impact on KPIs Objective function value (a), Total number of volunteers assigned (b), Number of volunteers assigned in groups (d), % of volunteer pairs with similar age IDs (e), % of volunteers assigned in groups with similar skill level values (f) but does not have a main effect impact on the KPIs total number of tasks with volunteer assignments (c) and % of total task-skill requirements met (g).

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	9487	9487	225.508	< 2e-16	***
max.menu.size	2	107310	53655	1275.443	< 2e-16	***
Pickiness	1	419423	419423	9970.186	< 2e-16	***
Heterogen	1	289356	289356	6878.329	< 2e-16	***
CL.diff	1	12179	12179	289.519	< 2e-16	***
Age.ID.diff	1	2100	2100	49.909	7.73E-12	***
min.g.size:max.menu.size	2	1628	814	19.351	9.93E-09	***
min.g.size:Pickiness	1	6114	6114	145.344	< 2e-16	***
max.menu.size:Pickiness	2	441	221	5.247	0.00565	**
min.g.size:Heterogen	1	1239	1239	29.45	1.02E-07	***
max.menu.size:Heterogen	2	1529	764	18.169	2.91E-08	***
Pickiness:Heterogen	1	291239	291239	6923.082	< 2e-16	***
min.g.size:CL.diff	1	1	1	0.033	0.85681	
max.menu.size:CL.diff	2	190	95	2.261	0.10561	
Pickiness:CL.diff	1	1020	1020	24.246	1.27E-06	***
Heterogen:CL.diff	1	770	770	18.301	2.39E-05	***
min.g.size:Age.ID.diff	1	0	0	0.011	0.91786	
max.menu.size:Age.ID.diff	2	124	62	1.479	0.22907	
Pickiness:Age.ID.diff	1	316	316	7.505	0.00644	**

Heterogen:Age.ID.diff	1	66	66	1.565	0.2117	
CL.diff:Age.ID.diff	1	0	0	0.004	0.94688	

Table A6: Normalized 'Objective Function' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	4	4	0.592	0.442133	
max.menu.size	2	17600	8800	1171.846	< 2e-16	***
Pickiness	1	81544	81544	10858.95	< 2e-16	***
Heterogen	1	51927	51927	6914.935	< 2e-16	***
CL.diff	1	1656	1656	220.548	< 2e-16	***
Age.ID.diff	1	482	482	64.126	1.44E-14	***
min.g.size:max.menu.size	2	35	18	2.349	0.096884	.
min.g.size:Pickiness	1	43	43	5.755	0.016926	*
max.menu.size:Pickiness	2	39	20	2.599	0.075663	.
min.g.size:Heterogen	1	16	16	2.136	0.144727	
max.menu.size:Heterogen	2	386	193	25.734	3.30E-11	***
Pickiness:Heterogen	1	52209	52209	6952.499	< 2e-16	***
min.g.size:CL.diff	1	25	25	3.288	0.070584	.
max.menu.size:CL.diff	2	182	91	12.12	7.90E-06	***
Pickiness:CL.diff	1	172	172	22.94	2.40E-06	***
Heterogen:CL.diff	1	114	114	15.219	0.000113	***
min.g.size:Age.ID.diff	1	8	8	1.05	0.306056	
max.menu.size:Age.ID.diff	2	58	29	3.851	0.022084	*
Pickiness:Age.ID.diff	1	71	71	9.498	0.002207	**
Heterogen:Age.ID.diff	1	23	23	3.036	0.08223	.
CL.diff:Age.ID.diff	1	17	17	2.213	0.137658	

Table A7: Normalized 'Total Number of Volunteers Assigned' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	3	3	2.166	0.1419	
max.menu.size	2	748	374	283.867	< 2e-16	***
Pickiness	1	6638	6638	5041.495	< 2e-16	***
Heterogen	1	5665	5665	4302.533	< 2e-16	***
CL.diff	1	2	2	1.153	0.2835	
Age.ID.diff	1	0	0	0.04	0.84245	
min.g.size:max.menu.size	2	9	5	3.5	0.03116	*
min.g.size:Pickiness	1	4	4	2.66	0.10373	
max.menu.size:Pickiness	2	159	80	60.456	< 2e-16	***
min.g.size:Heterogen	1	0	0	0.078	0.78082	
max.menu.size:Heterogen	2	35	18	13.335	2.52E-06	***

Pickiness:Heterogen	1	5748	5748	4365.381	< 2e-16	***
min.g.size:CL.diff	1	0	0	0.014	0.90507	
max.menu.size:CL.diff	2	0	0	0.019	0.98119	
Pickiness:CL.diff	1	5	5	4.116	0.04318	*
Heterogen:CL.diff	1	0	0	0.014	0.90507	
min.g.size:Age.ID.diff	1	0	0	0.04	0.84245	
max.menu.size:Age.ID.diff	2	0	0	0.101	0.90372	
Pickiness:Age.ID.diff	1	0	0	0.014	0.90507	
Heterogen:Age.ID.diff	1	0	0	0.002	0.96829	
CL.diff:Age.ID.diff	1	0	0	0.267	0.60538	

Table A8: Normalized 'Number of Tasks with Assignments' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	95	95	8.932	0.00298	**
max.menu.size	2	23680	11840	1109.633	< 2e-16	***
Pickiness	1	98843	98843	9263.335	< 2e-16	***
Heterogen	1	59844	59844	5608.462	< 2e-16	***
CL.diff	1	2762	2762	258.849	< 2e-16	***
Age.ID.diff	1	689	689	64.552	1.20E-14	***
min.g.size:max.menu.size	2	59	30	2.783	0.06312	.
min.g.size:Pickiness	1	93	93	8.754	0.00328	**
max.menu.size:Pickiness	2	368	184	17.241	6.81E-08	***
min.g.size:Heterogen	1	20	20	1.84	0.17572	
max.menu.size:Heterogen	2	1285	642	60.208	< 2e-16	***
Pickiness:Heterogen	1	58391	58391	5472.276	< 2e-16	***
min.g.size:CL.diff	1	116	116	10.844	0.00108	**
max.menu.size:CL.diff	2	269	135	12.618	4.95E-06	***
Pickiness:CL.diff	1	116	116	10.877	0.00107	**
Heterogen:CL.diff	1	166	166	15.546	9.58E-05	***
min.g.size:Age.ID.diff	1	34	34	3.199	0.07447	.
max.menu.size:Age.ID.diff	2	94	47	4.419	0.01267	*
Pickiness:Age.ID.diff	1	87	87	8.133	0.00458	**
Heterogen:Age.ID.diff	1	23	23	2.172	0.14139	
CL.diff:Age.ID.diff	1	29	29	2.746	0.09833	.

Table A9: Normalized 'Number of Volunteers Assigned in Groups' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	0.04	0.04	2.873	0.090934	.
max.menu.size	2	0.13	0.065	4.623	0.010394	*
Pickiness	1	0.264	0.264	18.698	1.97E-05	***

Heterogen	1	0.166	0.166	11.764	0.000672	***
CL.diff	1	0.035	0.035	2.514	0.113671	
Age.ID.diff	1	6.123	6.123	434.322	< 2e-16	***
min.g.size:max.menu.size	2	0.107	0.053	3.792	0.02342	*
min.g.size:Pickiness	1	0.015	0.015	1.076	0.300343	
max.menu.size:Pickiness	2	0.06	0.03	2.133	0.119928	
min.g.size:Heterogen	1	0.068	0.068	4.826	0.028654	*
max.menu.size:Heterogen	2	0.016	0.008	0.584	0.557904	
Pickiness:Heterogen	1	0.19	0.19	13.473	0.000277	***
min.g.size:CL.diff	1	0.005	0.005	0.356	0.550924	
max.menu.size:CL.diff	2	0.015	0.008	0.538	0.584372	
Pickiness:CL.diff	1	0.038	0.038	2.679	0.102514	
Heterogen:CL.diff	1	0.019	0.019	1.366	0.243311	
min.g.size:Age.ID.diff	1	0.004	0.004	0.281	0.596136	
max.menu.size:Age.ID.diff	2	0.194	0.097	6.889	0.001154	**
Pickiness:Age.ID.diff	1	0.411	0.411	29.175	1.18E-07	***
Heterogen:Age.ID.diff	1	0.277	0.277	19.654	1.22E-05	***
CL.diff:Age.ID.diff	1	0	0	0.008	0.928508	

Table A10: Normalized '% of volunteer pairs with similar age IDs' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	617	617	7.215	0.00755	**
max.menu.size	2	101654	50827	594.659	< 2e-16	***
Pickiness	1	468121	468121	5476.84	< 2e-16	***
Heterogen	1	216188	216188	2529.324	< 2e-16	***
CL.diff	1	3829	3829	44.799	7.88E-11	***
Age.ID.diff	1	4506	4506	52.714	2.21E-12	***
min.g.size:max.menu.size	2	334	167	1.953	0.14323	
min.g.size:Pickiness	1	1921	1921	22.475	3.02E-06	***
max.menu.size:Pickiness	2	1837	918	10.746	2.89E-05	***
min.g.size:Heterogen	1	550	550	6.437	0.01158	*
max.menu.size:Heterogen	2	5723	2861	33.476	4.09E-14	***
Pickiness:Heterogen	1	209412	209412	2450.047	< 2e-16	***
min.g.size:CL.diff	1	449	449	5.253	0.02246	*
max.menu.size:CL.diff	2	900	450	5.263	0.00557	**
Pickiness:CL.diff	1	3192	3192	37.351	2.45E-09	***
Heterogen:CL.diff	1	313	313	3.656	0.05662	.
min.g.size:Age.ID.diff	1	166	166	1.941	0.16438	
max.menu.size:Age.ID.diff	2	937	468	5.478	0.00451	**
Pickiness:Age.ID.diff	1	515	515	6.03	0.01451	*
Heterogen:Age.ID.diff	1	4	4	0.052	0.81998	
CL.diff:Age.ID.diff	1	237	237	2.776	0.09648	.

Table A11: Normalized '% of volunteers assigned in groups with similar skill levels' Value ANOVA results

Main and Interaction Factors	Degrees of Freedom	Sum of Squares	Mean of Squares	F value	p-value	Significance
min.g.size	1	0	0	0.135	0.713526	
max.menu.size	2	3.987	1.993	772.743	< 2e-16	***
Pickiness	1	13.844	13.844	5366.983	< 2e-16	***
Heterogen	1	10.763	10.763	4172.391	< 2e-16	***
CL.diff	1	0.085	0.085	32.808	2.08E-08	***
Age.ID.diff	1	0.005	0.005	1.747	0.187042	
min.g.size:max.menu.size	2	0.014	0.007	2.655	0.071619	.
min.g.size:Pickiness	1	0.055	0.055	21.448	5.01E-06	***
max.menu.size:Pickiness	2	0.082	0.041	15.931	2.28E-07	***
min.g.size:Heterogen	1	0.004	0.004	1.376	0.241532	
max.menu.size:Heterogen	2	0.048	0.024	9.301	0.000114	***
Pickiness:Heterogen	1	10.696	10.696	4146.735	< 2e-16	***
min.g.size:CL.diff	1	0.018	0.018	7.019	0.008404	**
max.menu.size:CL.diff	2	0.003	0.002	0.582	0.559477	
Pickiness:CL.diff	1	0.02	0.02	7.66	0.005925	**
Heterogen:CL.diff	1	0	0	0.002	0.965607	
min.g.size:Age.ID.diff	1	0	0	0.054	0.816267	
max.menu.size:Age.ID.diff	2	0.004	0.002	0.855	0.42601	
Pickiness:Age.ID.diff	1	0.001	0.001	0.27	0.603727	
Heterogen:Age.ID.diff	1	0	0	0.002	0.966333	
CL.diff:Age.ID.diff	1	0	0	0.035	0.851341	

Table A12: Normalized '% of total task-skill requirements met' Value ANOVA results

X8 - Data Generation for Design of Experiments

X8.1 - Sets:

- R is the set of all tasks, defined using the partner NPO's volunteer dataset and an understanding of the services they provide. 12 distinct tasks are defined, many of which are required for each NPO project simultaneously. Examples of current NPO projects are: XYZ Course (YouTube Channel), XYZ Video Series (YouTube Channel), XYZ Podcasts, XYZ Pamphlets. For experiments, Set R has a total of 24 tasks representing tasks related to two projects of partner NPO.
- V is the set of all volunteers and is determined based on the total available volunteer-to-task assignment slots possible for the fixed value of maximum group size ($\check{g}_r = 3$) used. Set V has 24 x 3 = 72 volunteers.
- S is the set of all skills (S1-S10)(Appendix X2). Some combination of skills is required for each task (Appendix X1). Definition of task scope and skill requirements is obtained from understanding of the partner NPO's tasks and their requirements.

- Q is the set of all scenarios, we run design of experiments with 10 scenarios (σ) per replication of an experiment.
- Replication of experiment refers to the experiment runs where input parameters such as volunteer information is unique per replication. We run 5 replications per experiment run for a particular factor and level.

X8.2 - Input Parameters:

- H_v = age ID of volunteer $v \in V$, generated using the discrete probability distribution of volunteers in each age ID category (1-10), obtained from NPO's volunteer dataset (Figure 3).
- i_v = number of task preference selections made by a volunteer in their onboarding form entry. Volunteers selecting 5 task preferences are more dedicated and eager to volunteer versus volunteers who select only 1 task of 5 options offered (values 1-5). Generated using discrete probability distribution of the number of task preference selections volunteers made on their onboarding form, obtained from NPO's volunteer dataset (Figure 5).
- j_{rv} = probability of a volunteer selecting a task $r \in R$, given the age ID of the volunteer is H_v, (values between 0-1). Generated using conditional probability distributions of task selections for each age ID category, obtained from NPO's volunteer dataset (Figure 4).
- E_{rs} = skill level required in skill $s \in S$, on task $r \in R$, for successful completion, (values 0-10)
    - Data for E_{rs} is designed using detailed information available on steps required to complete different NPO tasks. An educated estimate of the skills required on tasks and the respective skill levels required is made.
    - Values remain the same across all experiments and are determined by the NPO ahead of time when establishing projects and their associated tasks to be completed. The E_{rs} values used for our computational experiments can be found in Appendix X1.

- C_{vs} = skill level of volunteer $v \in V$ in skill $s \in S$, (values 1-7). Generated using the distribution of volunteer skills obtained from the NPO's volunteer dataset (Figure 6). A representative set of volunteers is generated using distribution information for our computational experiments.

X9 - MC Methodology Factor Levels – Model performance results for varying environmental instances when menu size b = 3 vs. menu size b = 5 are created

Menu Size b = 3 Created vs. Menu Size b = 5 Created								
b = 3 vs. 5	Avg Over All Env. Instances	5.39	6.85	2.86	8.46	4.00	4.43	4.45
	p-value	<<0.05	<<0.05	<<0.05	<<0.05	0.25	0.68	0.65
	LPV*	2.68	5.31	0.11	6.57	2.53	1.34	3.10
	HPV*	9.50	10.97	6.82	14.12	5.26	9.83	7.84
	NH*	5.84	5.95	6.65	5.69	5.95	4.48	2.54
	SH*	5.10	6.64	0.21	9.14	2.06	3.59	4.69

*LPV = Low Picky Volunteers, HPV = High Picky Volunteers, NH = No Heterogeneity, SH= Some Heterogeneity

The % difference between implementing factor-levels of the MC methodology, for different environmental factors, for each KPI (a) – (g)

Looking at the average percent change and the impact on each KPI between implementing a menu size of 3 versus menu size 5, it is observed that there is a statistically significant increase in KPIs (i) – (vii), but KPIs % of volunteer pairs with similar age IDs (v), % of volunteers assigned in groups with similar skill level values (vi), and % of total task-skill requirements met (vii) do not see a statistically significant increase. These KPIs represent the percent of homogenous groups formed and percent of total skill requirements met, which are already being improved significantly when a menu of size 3 is implemented.

Figure A1 displays the improvements achieved in all KPIs upon implementing a larger menu size of 5, for the environment factor 'high picky volunteers' (HPV).

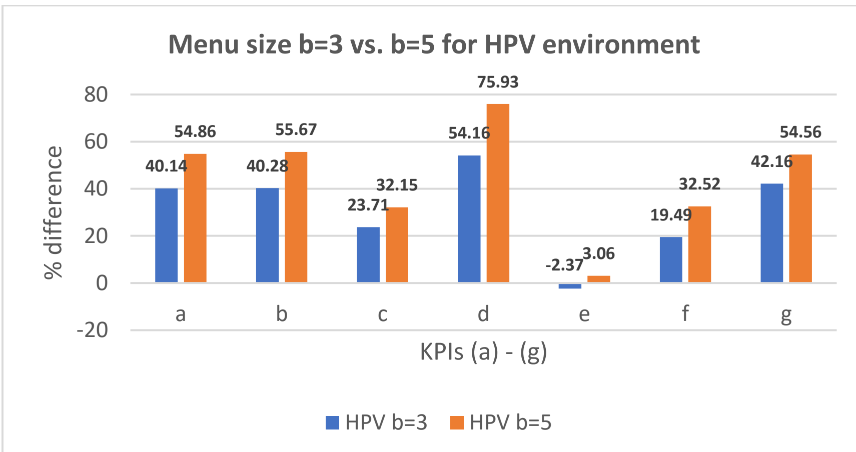


Figure A1: Impact of implementing menu size b=3 vs. b=5 for environments with High Picky Volunteers (HPV)