MODELING MARGINALIZATION:
EMERGENCE, SOCIAL PHYSICS, AND SOCIAL ETHICS OF BULLYING

Themis D. Xanthopoulou
Andreas Prinz

Department of Information and Communication Technology
University of Agder
Jon Lilletuns vei 9
Grimstad, Aust-Agder, Norway
themis.d.xanthopoulou@uia.no
andreas.prinz@uia.no

Ivan Puga-Gonzalez
F.LeRon Shults

Department of Global Development and Planning
University of Agder
Universitetsveien 19
Kristiansand, Vest-Agder
ivan.puga-gonzalez@uia.no
leron.shults@uia.no

ABSTRACT

In this paper, we outline the construction and initial simulation experiment results of the Marginalization model (MARG). We experiment under different group parameters because the theoretical paradigm we follow views bullying as a result of social processes. Our primary research question explores the possibility of bullying emergence as agents select interaction partners in a university setting. Based on the simulated process, our results take indications of the stress of marginalization in a student group as a proxy for emergent marginalization. MARG simulates two types of interactions between pairs of students: forced and hang-out interactions. In the latter, students decide whether to interact based on individual preferences formed by social norms and individual tolerance related to those norms. The emergence of intensified marginalization from MARG processes leads to some ethical considerations and provides ground for discussions concerning suitable interventions.

Keywords: marginalization, bullying, social simulation, ethics

1 INTRODUCTION

A recent report by Ingrid Lund found that 9% of students at universities and colleges in Norway have experienced bullying (Lund 2017). Lund's report attracted a lot of attention, and a working group at the University of Agder was founded by the rector to address the issue. The strong reaction comes as no surprise since bullying is linked with mental health problems and violence (Blood and Blood 2016, Wolke and Lereya 2015). In addition to health issues, students who are bullied are more likely to drop out of school (Cornell et al. 2013). Notwithstanding the ever-growing body of research on bullying, there is no concrete explanation of how bullying emerges (Turner et al. 2015). Most people would agree that it is wrong and unethical to bully someone. But is it possible for this "unethical" act to emerge from an ethically acceptable action such as choosing an interaction partner? In this paper, we explore in an agent-based computer model the effect of University students' partner choices on bullying.

2 THEORIES AND METHODS OF BULLYING

Bullying is a concept employed to describe a wide range of behaviors (Cohen and Brooks 2018), and currently, there is no consensus among researchers for the definition (Turner et al. 2015). The large number of interpretations make it very difficult to impossible to integrate all definitions into one model consistently. In addition to the myriad of behaviors, there are currently two paradigms for understanding bullying (Schott and Søndergaard 2014). The first and older one focuses on the individual traits of involved parties and the second one on existing social dynamics as the force that drives bullying. The functional difference between the two is that the first paradigm places responsibility on the individual while the second one on the interplay of the dynamics. Several studies support the idea that there are more than individual traits that play a role in bullying. Salmivalli's work highlighted the group character of bullying by introducing the roles of bystanders, defenders of the victim, and supporters of the bully (Salmivalli et al. 1996). In addition, Paluck et al.' s study indicated the importance of changing norms to reduce conflict in schools (Paluck, Shepherd, and Aronow 2016). Finally, the emergence of a socio-ecological model where different norms relate to the bullying phenomenon supports the adoption of the second paradigm to explain bullying (Espelage and Swearer 2010). Based on these findings, we approach bullying with the second paradigm. Based on a popular definition within this paradigm, we operationalize bullying as the "intensification of the processes of marginalization that occur in the context of dynamics of inclusion-exclusion, which shape groups. Bullying happens when physical, social, or symbolic exclusion becomes extreme, regardless of whether such exclusion is experienced or intended" (Schott and Søndergaard 2014).

The context of our paper is bullying at the University. To be more specific, we are interested in the exclusion of students from interactions that take place in their free time but within the space of the University. Some examples would be interactions during recess and group study. Exclusion is another ambiguous concept (Peace 2001) with multiple dimensions (Mathieson et al. 2008), and we need to define it further. We focus on exclusion related to communication with peers in the educational environment. We acknowledge three types of interactions: positive, negative, and refused interactions. We view instances of refused and negative interactions as instances of exclusion. For our paper, we consider marginalization as the culmination of exclusion events. Marginalization is expressed for an individual when she has a large ratio of negative or refused interactions or when, on average, the interaction partners are not attracted to the individual and do not want to interact. We define the "intensified" component by using thresholds.

We want to understand if a process such as the selection of interaction partners, which is not considered unethical, could lead to bullying, as defined in the previous paragraph. Based on the theory developed by Thibault and Kelly, the result of the interactions is based on the compatibility of each interaction partner's characteristics (Thibaut and Kelley 1959). Compatibility implies the existence of a sort of evaluation in the characteristics of the interaction partner. People choose interaction partners based on the outcomes of previous interactions. In our paper, we view compatibility as a matter of individual preference. Different cultures favor different types of behaviors (Hofstede, Hofstede, and Minkov 2010), and individual preferences are formed within these contexts. We acknowledge that our approach is limited in tracking the manifested bullying in a case study because it includes only a related process and that it will most probably only identify a tendency towards bullying. On the other hand, the limitations do not prevent us from reflecting ethically on the emergence of bullying.

We chose the Agent-Based Modeling (ABM) methodology to explore bullying dynamics in relation to the selection of interaction partners. ABMs have been used before to address bullying but mostly based on predefined roles in bullying for the agents and individual perspectives. Therefore, previous work links to the framework of the first paradigm for explaining bullying. Tseng et al. created an ABM model of bullying using Social Impact Theory and defined student roles as the victim, the bully, and the bystanders (Tseng et al. 2014). Thawiworadilok et al. used an ABM based on game theory (prisoner dilemma game) and assigned with probability student roles to test the effect of the victim's compliance with the bully (Thawiworadilok, Songhori, and Terano 2017).

In contrast, we focus on the second paradigm and shift the focus onto the group perspective. Closer to the second paradigm with regards to non-predefined roles, with results still on the individual level, is the model by (Maeda, Anezaki, and Takeuchi 2006). They identify students with a set of values to portray their interests and simulate homogenization processes to understand which students will get excluded and thus have higher chances of being bullied. They showed that students with rare values or interests are more likely to be left out in interactions. In our model, we do not include homogenization processes because we consider them a defense mechanism to counteract intense marginalization. Our main question is, "Can our model produce intensified marginalization as a result of the social process of selecting interaction partners based on individual preferences formed by social norms?"

3 MARG: THE BULLYING MODEL

3.1 Model Overview

MARG is an agent-based model written in Netlogo version 6.1.1 (available here: https://github.com/themisdx/MARG https://github.com/themisdx/MARG) designed to represent the interactions that occur in a university setting among a group of students (the number of students is a model variable called "num-students"), without previous knowledge of each other. Each time step represents a 'day' and therefore the simulation last 100 number of time steps, resembling thus a whole university semester. The most important variables are shown in Table 1 and Table 2. At each time step, agents have two types of interactions: classroom interactions and hang-out interactions. All interactions are dyadic. During classroom interactions, agents cannot refuse to start an interaction with others; therefore, we call these forced interactions. In contrast, during hang-out interactions, agents may refuse to start an interaction with another agent, and thus we call these free interactions. At each time step and for each type of interaction, agents are paired N times in randomly selected dyads, with N = (num-students – 1)/2.

Interactions result in positive or negative evaluations. These evaluations are based on the comparison between the interaction partner's characteristics and 'ideal' values (see below for details). For every other agent, agents have a variable called "attraction" that represents their likeness to the other agent and thus their disposition to interact with it. This variable is unidirectional, which means that two agents may have a dissimilar attraction towards each other. When the value is 1, an agent feels maximum attraction for that specific agent, and when it is 0, the agent feels minimum attraction.

An important element of hang-out interactions is the element of choice. During hang-out interactions (free interactions), after an agent has been randomly paired with another, it first decides whether it wants to interact or not with its partner by comparing its attraction link towards the other agent with a random number between 0 and 1. If the attraction is higher than the random number, the agent proceeds with the interaction (note that during forced -classroom- interactions, the first agent always starts an interaction regardless of the value of the attraction link towards its partner). Then, in the hang-out interactions, the other agent must decide whether to proceed with the interaction on the basis of the same decision rule. In the hang-out interactions, if the first agent refuses, the second one is unaware of the decision. On the other hand, when the first agent starts the interaction and the interaction partner rejects the interaction, the first agent will acknowledge this as a refused interaction and will increment its #refusedinteractions by 1. In both forced and free interactions, if an interaction is positive, then the agent increases its attraction towards the interaction partner by the amount of "Attraction_change" and thus the likelihood of interacting with that agent in the future. If, on the other hand, the interaction is negative, then the agent decreases its attraction link by the same amount and thus the likelihood of interacting with that agent in the future.

Table 1: Simulation Parameters.

Group variables	**Description**	**Range**	**Increment**
Num-students	Number of students in the simulation	10-50	10
Num-internal-characteristics	Number of internal characteristics per agent	Fixed: 10	NA
Num-external-characteristics	Number of external characteristics per agent	Fixed: 10	NA
Average_char	Average value "Average_char" and standard deviation "Stdev_char" of the truncated (min 0 and max 1) normal distribution from which values are drawn for the agents' characteristics	Fixed: 0.5	NA
Stdev_char		0.1-0.5	0.1
Attraction_change	Change in the attraction link of an interaction partner after a positive or negative interaction.	0.02-0.1	0.02
Attitude	Initial attraction link value towards the other interaction partners.	Fixed: 0.5	NA
Max_judg	Maximum value of the uniform distribution from which values of tolerance are drawn.	0.1-0.5	0.1
Charlearned	Number of internal-characteristics an agent learns about an interaction partner during an interaction	Fixed: 1	NA
In_ideal_chars	Optimal cultural value of the internal characteristics (shared for all agents).	Scenario 1: 0.5 Scenario 2: 1.0	NA NA
Ex_ideal_chars	Optimal cultural value of the external characteristics (shared for all agents).	Scenario 1: 0.5 Scenario 2: 1.0	NA NA

At initialization, Attraction towards others is set to the value of "Attitude". In our simulations, Attitude equals to 0.5, so that all agents have the same chance of accepting/rejecting an interacting. Further, the decision rule ensures that agents will have a chance to interact with those they will consider not so attractive and reject interactions from those whom they will feel a strong attraction to.

3.2 Characteristics of Agents and Learning

Each agent is endowed with a set of external and internal characteristics. "External-characteristics" represent conspicuous features of persons such as physical appearance; "internal-characteristics" represent more personal features such as personal interests, way of thinking, etc. which are only known through communication. "Num-internal-characteristics" and "num-external-characteristics" allow the manipulation of the number of characteristics in the model. The values of the external and internal characteristics among agents are heterogeneous. Besides these characteristics, agents are also provided with an 'ideal' set of external and internal values. The values are controlled by the variables "In_ideal_chars" and "Ex_ideal_chars" This set of 'ideal' values is the same among all agents and represents the 'cultural' traits, affected by cultural norms, agents look for when interacting with others (see Table 1). The closer the values of the extrinsic and intrinsic characteristics of an agent to the ideal values, the higher the likelihood of being liked by others. However, cultural norms are not the only factor in evaluating the interaction partner positively. Each agent is assigned its own "tolerance" variable. The combination of the ideal and the tolerance value form the individual preference (see Table 2). The higher the tolerance, the wider the range of the individual preference, the less the effect of norms in liking another agent.

At the beginning of the simulations, the agents only know the external characteristics of others; internal characteristics are learned during interactions. The number of internal-characteristics learned during each interaction is expressed by the variable "char-learned", and the characteristics themselves each agent knows for one other agent are stored in an array called "known-indices".

3.3 Evaluations during Interactions

Whether an interaction is positive or negative depends on the evaluation of the agent's characteristics values (extrinsic and intrinsic) against those of the 'ideal' set of characteristics. If the value of the agent's characteristic falls within the individual preference of the evaluator, then the evaluation is positive (value 1), and otherwise, it is negative (value -1). As a result, the higher the tolerance of an agent, the more likely it is to evaluate another agent positively and vice versa. The agent evaluates all extrinsic characteristics and known intrinsic characteristics of the interaction partner. If the sum of these evaluations is positive, then the interaction is positive, and the agent will register +1 #positiveinteractions. In the opposite case, the agent will register +1 #negativeinteractions. The consequence of learning more characteristics about the interaction partners is that the evaluations change as time progresses.

3.4 Experimental Set-Up

Table 1 shows the parameter space for our simulation. All simulation parameters are group parameters due to the adoption of the second paradigm described in Section 2. For the experiments, we manipulate the group parameters, and they form the agent parameters shown in Table 2. The method of generation of the agent parameters is also described in Table 2. We ran simulations under two different scenarios varying four parameters (shown in Table 1). For Scenario 1 or 2 and a combination of parameter set (n=625 or 5^4), we ran five replications (to ensure the reliability of the results), leading to a total of 6250 simulations. Further, based on the values of Max_jud and Stdev_char (Table 1), we categorize simulations into low/medium/high tolerance and diversity, as shown in Table 3.

3.5 Data Collection, Bullying Metrics, and Data Analysis

For each agent, we collected the variables #positiveinteractions, #negativeinteractions, #refusedinteractions, and the average attraction value the other agents felt towards it (calculated in variable "average_attraction_in") at the end of the simulations (see Table 2). With the data collected and taking the perspective of the second paradigm, which relates bullying to social dynamics, we measured bullying as the intensified marginalization 'experienced' by an agent calculated in two different ways:

Metric 1. *"Marginalized?" was based on the exclusion index. When the index passes the threshold of 0.8, meaning that more than 80% of the interactions of the agent are negative, then Exclusion? was set to true. The logic behind the exclusion index is to capture the number of exclusion experiences perceived by the* *agent in relation to all its interactions. The threshold represents the intensified marginalization. The exclusion index is shown in equation (1):*

$$\text{Exclusion index} = \frac{\#NegativeInteractions + \#RefusedInteractions}{\#NegativeInteractions + \#RefusedInteractions + \#PositiveInteractions}. \quad (1)$$

Metric 2. *"Marginalized_attraction?" was based on the average_attraction_in: when the value of this metric is below 0.2, meaning that on average other agents reject an interaction 8 out of 10 times with the target agent, then Marginalized_attraction? was set to true. This metric represents the exclusion received from others. The threshold represents the intensified marginalization.*

Table 2: Agents' variables and Data Collection variables.

Agent Variables	**Description**
Tolerance	Individual value assigned to each agent from a uniform distribution with min = 0 and max = Max_jud
Internal-characteristics	The intrinsic characteristics of the agent. Values range between 0 and 1 and are drawn from a normal distribution with average = Average_char and standard deviation = Stdev_char
External-characteristics	The extrinsic characteristics of the agent. Values range between 0 and 1 and are drawn from a normal distribution with average = Average_char and standard deviation = Stdev_char
Individual preference	Range of preference of a characteristic = ideal value ± tolerance
Attraction	A variable representing how much an agent 'likes' another one; values range from 0 (low attraction) to 1 (high attraction). Each agent has one parameter value for each other agent.
known-indices	The exact internal-characteristics an agent knows from another one at a given time step in the simulation for each agent. Each agent has one array of values for each other agent.
Data Collection Variables	**Description**
#positiveinteractions	The total positive interactions the agent had in the simulation
#negativeinteractions	The total negative interactions the agent had in the simulation
#refusedinteractions	The total refused interactions the agent had in the simulation
Average_attraction_in	The average value of incoming attraction of all other agents towards this agent
Students_ marginalized	The percentage of agents excluded in a specific simulation based on a threshold value and the agents' exclusion index

Students_ marginalized_attraction	The percentage of agents excluded in a specific simulation based on a threshold value and the agents' average_attraction_in

These metrics represent two different aspects of exclusion that do not necessarily coincide. Metric 1 represents how a specific agent perceives exclusion from others, and metric 2 represents the perception of the group (average attraction) towards the specific agent. Hence, there may be cases in which a highly tolerant agent perceives positively most of its interactions (because it evaluates others positively), yet itself is not liked and gets often refused interactions from others. In this case, metric 1 will not indicate a lot of negative interactions, while metric 2 will a marginalization signal. On the contrary, it can happen that an intolerant agent dislikes everybody and isolates itself (by refusing to interact with others) even though it is liked by most others. In that case, metric 1 will indicate marginalization, while metric 2 will not give this signal.

Even though we performed data collection on the individual parameters, we scaled the data to the group level so that we can have conclusions for the social dynamics. For this reason, we calculated the percentage of marginalized agents at the group level for each simulation separately based either on metric 1 or 2, these variables were named "Students_ marginalized" and "Students_ marginalized _attraction", respectively.

Table 3: Categorization of Simulations according to the values of Max_jud and Stdev_char.

Simulation type	**Variable value**
Low Tolerance	Max_judg = 0.1
Medium Tolerance	Max_judg = 0.2 or 0.3
High Tolerance	Max_judg = 0.4 or 0.5
Low Diversity	Stdev_char = 0.1
Medium Diversity	Stdev_char = 0.2 or 0.3
High Diversity	Stdev_char = 0.4 or 0.5

4 FINDINGS: THE SOCIAL PHYSICS OF BULLYING

We borrow the term "social physics" from (Pentland 2014) to indicate that we study the dynamics and mechanisms by which ideas, behaviors, and interactions emerge and spread within human populations. The initial simulation experiments with MARG shed light on our research question. Pearson correlations between the group exclusion based on the two different metrics ("Students_ marginalized" and "Students_ marginalized_attraction") showed that these variables were highly correlated: $r=0.94$ $P \leq 0.001$ for Scenario 1 and $r=0.9$ $P \leq 0.001$ for Scenario 2. Due to the high correlation, we decided to proceed with the presentation of results for "Students_ marginalized", which is based on Metric 1. For Metric 1, we calculated the descriptive statistics for all simulations of Scenario 1 (see Table 4) and 2 (results in the text). Due to the high variability of the results in Scenario 1, we performed a linear model analysis (see Table 5) and mixed model analysis. The linear model displayed a better explanation for the variance of the results.

In scenario 2, where the idealized values are all set to 1, the values of the agents' characteristics (internal and external) will seldom coincide because agents' characteristics have a mean value of 0.5. Hence, during the evaluation of interactions, most agents will have a negative experience with all others resulting in a high exclusion index. Therefore, agents are driven towards marginalization. Nevertheless, when tolerance is high, the percentage of marginalized students decreases because high tolerance broadens the agents' preferences, and thus, they find other partners attractive even though they have characteristic values far from the 'ideal' ones. This percentage of marginalized students is even lower when combined with higher diversity in the values of agents' characteristics because a high diversity in characteristics increases the chance that some agents have values close to the "ideal" ones. The descriptive statistics of Metric 1 for Scenario 2 showed 100% of students marginalized per simulation except for the cases where we have simulations characterized by high tolerance. When tolerance is high, medium, or low combined with high diversity, we have, on average, 83% of students marginalized per simulation, 93% of students marginalized per simulation, and 97% of students marginalized per simulation, respectively.

The results are more diverse when we look at Scenario 1 in which the idealized values coincide with the average value of the agents' characteristics (Table 4). As a result, the more a student's characteristics match the average values, the more likely it is for this agent to be included and vice versa. As in scenario 2, high tolerance reduces group marginalization because it broadens individual preferences. However, in contrast to Scenario 2, diversity has the opposite effect. The combination that favors marginalization is high diversity and low tolerance (100% of students marginalized per simulation) and medium diversity and low tolerance (98% of students marginalized per simulation) (Table 4).

On the other hand, low diversity favors low rates of exclusion (high rates of inclusion). Low diversity and high tolerance yield an average of 10% of students marginalized per simulation and low diversity and medium tolerance an average of 20% of students marginalized per simulation (Table 4). Interestingly, even though we see conditions leading to 100% of students marginalized per simulation, we do not see the opposite, which is simulations with 0% marginalization. The lowest average marginalization was 10% under the conditions of high Tolerance and low Diversity. It seems, thus, that in our model, the marginalization is more easily achieved than inclusion.

Table 4: The table shows the average percentage of marginalized students (based on Metric 1) per simulation for all simulations under specific conditions of tolerance and diversity. The results are for the simulation of Scenario 1 (total of 3125 simulations). The parentheses show the 1st and 3rd quartile for the specified result.

	Low Tolerance	**Medium Tolerance**	**High Tolerance**
Low Diversity	49% (40%-55%)	20% (14%-26%)	10% (5%-15%)
Medium Diversity	98% (100%-100%)	50% (35%-62%)	26% (20%-34%)
High Diversity	100% (100%-100%)	87% (75%-100%)	50% (40%-60%)

To better understand the effect of the input parameters (i.e. stdev_char, the max_judg, the attraction_change, and the num-students, Table 1) on the percentage of students marginalized per simulation (Metric 1), we performed a linear model analysis. We first ensured that parameters were not collinear and that the assumptions of linearity, homoscedastic, and normality of residuals are met. Results of the linear model analysis (using the lm function in R) are shown in table 5. The model explains a significant amount of the variance (Multiple R^2: 0.83, Adjusted R^2: 0.83). The input variables, Stdev_char (the diversity of agents' characteristics) and Max_judg (tolerance) parameters had a significant ($p < 0.001$), opposite but similar effect in size. None of the other input parameters had a significant effect on the percentage of marginalized students.

Max_judg and stdev_char have antagonistic effects: as the diversity among the values of the agents' characteristic increases, the percentage of students marginalized per simulation increases, and as the value of tolerance increases, the percentage of students marginalized per simulation decreases (Table 5). The effect is also visible in Table 4, where the diagonal combinations (low diversity-low tolerance, medium diversity – medium tolerance, high diversity – high tolerance) display similar simulation outcomes on average. Diversity increases marginalization because of the higher the diversity, the higher the deviation of the agents' characteristics from 0.5 (the idealized value in scenario 1). Thus, the more likely that agents evaluate interactions as negative. Note that this result is in contrast with scenario 2, where idealized values are set to 1. In scenario 2, a higher diversity decreases exclusion incidents since the higher the diversity, the higher the chance that agents' characteristics reach the idealized value of 1.

Table 5: Results of the Linear Model for Metric 1 and Scenario 1.

Linear Model			
Metric	**Effect**	**Standard error**	**P-value**
Intercept	0.58	0.01	<0.001
Stdev_char	1.41	0.02	<0.001
Max_judg	-1.53	0.02	<0.001
Attraction_change	-0.02	0.08	0.84
Num_students	<0.01	<0.01	0.69
Adjusted R^2: 0.83			

The limitations of the model include:

- The evaluation of each characteristic is polarized to either positive or negative, which leads to more extreme total evaluations. In addition, all characteristics are valued the same, whereas, in reality, some characteristics may be more important than others.
- We assumed ten external and ten internal characteristics per agent, and the effect of a different number of characteristics was not explored.
- All students share the same ideal values: a fact which implies the effect of one culture. Moreover, we have set the same ideal cultural value for all internal characteristics and the same ideal cultural value for all external characteristics.

In future work, we plan to look further into the aggregation of individual marginalization and the relationship of individual characteristics to the overall marginalization. Finally, we want to perform an extensive network analysis to track group formations among the students.

5 DISCUSSION: THE SOCIAL ETHICS OF BULLYING

With MARG, we were able to observe the emergence of intensified marginalization (defined with the use of thresholds and the metrics) for multiple parameter setups. The emergence occurred from everyday interactions at the university using one process from the context of inclusion-exclusion, the process of selecting interaction partners. Since we did not include more processes, we can argue that MARG simulates the societal stress of marginalization in students and not the actually manifested marginalization. The relation of the simulated vs. actual exclusion events cannot be estimated without further analysis. Nonetheless, the results raise questions about the "social ethics" of bullying. The latter term is highly contentious, as are the terms commonly related to it, such as exclusion and marginalization. However, generally speaking, most people find something "immoral" or at least morally problematic in the behaviors to which such terms commonly refer. On the other hand, most people do not find it morally problematic that individuals (e.g., university students) like to hang out with people whom they admire. MARG simulations showed that the latter could all too easily lead to the former.

Our modeling efforts here are part of a growing movement in the modeling & simulation profession toward surfacing the ethical assumptions and implications in model formulation, construction, and execution (Shults and Wildman 2019, Shults, Wildman, and Dignum 2018). In this context, we want to highlight the need to focus on *social* ethics as part of this ongoing conversation. All too often, ethical debates attend only or at least primarily to *individual* ethics; e.g., what rule should an individual follow, or what goals should the individual set? However, as models like MARG demonstrate, the social physics of human life is non-linear, and the emergent effects of goal setting are often far from what individuals intend. By constructing models that are attentive to social-ethical concerns in collaboration with philosophical ethicists and other stakeholders, computer scientists can contribute new tools and insights to conversations about how to mitigate the deleterious effects of social marginalization.

Within the boundaries of our model's limitations, we can draw conclusions on the social physics of bullying. Our findings are similar to those of the Schelling model that demonstrated how small shifts in preferring similar neighbors quickly lead to increased segregation between simulated agents (Schelling 1971). Unlike the Schelling model, however, MARG provides more insights about the drivers of the issue. The most notable conclusions from our simulations were the positive effect of tolerance on inclusion and the mixed effect of diversity on marginalization. The former is easily explained by the fact that tolerance determines individual preference, and a broad individual preference will result in a positive evaluation of the interaction partner and enhance inclusion. Explaining the latter requires more reflection. We would expect diversity to enhance the chance of students being included. Since we are simulating a monoculture with stabilized norms, diversity has a positive impact on inclusion when the culture idealizes characteristics outside of the average range. This is because, in our context, diversity introduces agents with characteristics closer to the ideal values. On the other hand, diversity worsens marginalization when a culture idealizes the average student because it increases the chance of deviation from the ideal value.

Based on our results, teaching tolerance is a universal strategy to combat bullying. Unfortunately, integrating tolerance into a person's thinking takes time, and therefore it can be considered a long-term strategy. On the other hand, creating groups with a preferred level of diversity is a short-term strategy, but it holds a risk. To decide about the diversity of a group requires a deep understanding of how ideal values relate to average characteristics in the group and knowledge of student characteristics. As it can be very difficult and ethically problematic to assess individuals' characteristics and store data about them, this short-term strategy applies only to specific groups such as small classes with consistent attendance. Overall, apart from focusing on altering the individuals' level of tolerance, educators and other stakeholders might also pay more attention to the function played by excessively high micro-level ideals in generating macro-level marginalization.

ACKNOWLEDGMENTS

The authors are grateful to the University of Agder for a grant to support the "Computer Tools for Modeling Social Conflict" project, which provides research funding for Themis Dimitra Xanthopoulou and Ivan Puga-Gonzalez.

REFERENCES

Blood, G. W., and I. M. Blood. 2016. "Long-term Consequences of Childhood Bullying in Adults who Stutter: Social Anxiety, Fear of Negative Evaluation, Self-Esteem, and Satisfaction with Life." *Journal of Fluency Disorders* vol. 50:pp. 72-84.

Cohen, J. W., and R. A. Brooks. 2018. "The Construction of School Bullying as a Social Problem." In *The Cambridge Handbook of Social Problems*, edited by A. J. Treviño, pp. 237-252. Cambridge:cCambridge University Press.

Cornell, D., A. Gregory, F. Huang, and X. Fan. 2013. "Perceived Prevalence of Teasing and Bullying Predicts High School Dropout Rates." *Journal of Educational Psychology* vol. 105 (1):pp. 138-149.

Espelage, D. L., and S. M. Swearer. 2010. "A Social-ecological Model for Bullying Prevention and Intervention: Understanding the Impact of Adults in the Social Ecology of Youngsters." In *Handbook of Bullying in Schools: An International Perspective.*, edited by S. R. Jimerson, S. M. Swearer and D. L. Espelage, pp. 61-72. New York, NY, US: Routledge/Taylor & Francis Group.

Lund, I. 2017. Mobbing i høyere utdanning, fleip eller fakta. https://www.universell.no/.

Mathieson, J., J. Popay, E. Enoch, S. Escorel, M. Hernandez, H. Johnston, and L. Rispel. 2008. "Social Exclusion Meaning, Measurement and Experience and Links to Health Inequalities." *A review of Literature. WHO Social Exclusion Knowledge Network Background Paper* vol. 1:pp. 91.

Paluck, E. L., H. Shepherd, and P. M. Aronow. 2016. "Changing Climates of Conflict: A Social Network Experiment in 56 Schools." *Proceedings of the National Academy of Sciences* vol. 113 (3):pp. 566-571. doi: 10.1073/pnas.1514483113.

Peace, R. 2001. "Social Exclusion: A Concept in Need of Definition?" *Social Policy Journal of New Zealand* (16):pp. 17-36.

Pentland, A. 2014. Social *Physics : How Good Ideas Spread the Lessons from a New Science*. New York: Penguin.

Salmivalli, C., K. Lagerspetz, K. Björkqvist, K. Österman, and A. Kaukiainen. 1996. "Bullying as a Group Process: Participant Roles and their Relations to Social Status within the Group." *Aggressive Behavior* vol. 22 (1):pp. 1-15.

Schelling, T. C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* vol. 1 (2):pp. 143-186.

Schott, R. M., and D. M. Søndergaard, eds. 2014. *School Bullying: New Theories in Context*. Cambridge: Cambridge University Press.

Shults, F. L., and W. J. Wildman. 2019. "Ethics, Computer Simulation, and the Future of Humanity." In *Human Simulation:* Perspectives*, Insights, and Applications*, edited by Diallo S. Y., Wildman W. J., Shults F. L. and Tolk A., pp. 21-40. Springer.

Shults, F. L., W. J. Wildman, and V. Dignum. 2018. "The Ethics of Computer Modeling and Simulation." 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden.

Turner, H. A., D. Finkelhor, A. Shattuck, S. Hamby, and K. Mitchell. 2015. "Beyond Bullying: Aggravating Elements of Peer Victimization Episodes." *School Psychology Quarterly* vol. 30 (3):pp. 366-384. doi: 10.1037/spq0000058.

Wolke, D., and S. T. Lereya. 2015. "Long-term Effects of Bullying." *Archives of Disease in Childhood* vol. 100 (9):pp. 879-885.

AUTHOR BIOGRAPHIES

THEMIS XANTHOPOULOU is a doctoral candidate at the University of Agder currently working on computer tools to address bullying. She is passionate about interdisciplinary research and draws knowledge from diverse scientific fields. Her email address is themis.d.xanthopoulou@uia.no.

IVAN PUGA-GONZALEZ is a post-doctoral research fellow at the University of Agder. His research interest is on the self-organization of complex social systems. More specifically, on how complexity at the group level may emerge from local interactions and simple heuristics at the individual level. He is particularly interested in human and non-human primates' social systems. He holds a Ph.D. in Theoretical Biology from the University of Groningen. His email address is ivan.puga-gonzalez@uia.no.

F. LERON SHULTS is Professor at the Institute for Global Development and Planning at the University of Agder in Kristiansand, Norway, and Scientific Director of the Center for Modeling Social Systems at NORCE. He has published 18 books and over 120 scientific articles on topics within a variety of fields, including social science, philosophy, cognitive science of religion, and computer science. His email address is leron.shults@uia.no.

ANDREAS PRINZ is Professor for systems development at the department of information and communication technology at the University of Agder, Campus Grimstad. He is interested in systems modeling and formal languages to express systems. He has experience in modeling systems, both technical and social, and in validating and verifying system models. His email address is andreas.prinz@uia.no.