

TERM ESSAY ON
“DOES INTEGRATION CHANGE GENDER ATTITUDES? THE
EFFECT OF RANDOMLY ASSIGNING WOMEN TO TRADITIONALLY
MALE TEAMS” DAHL ET AL. (2021)

1. Summary:

The main research question in the article “Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams” Dahl et. al. (2021) is if the integration of women into traditionally male-dominated environments can cause changes in men’s gender attitudes towards women.

The researchers conducted a field experiment in the Norwegian military, in which some men were randomly assigned to work with women during an eight-week initialisation boot camp. The empirical findings of this study conclude that integration with women caused men to become more egalitarian in their attitudes toward mixed-gender productivity, gender roles, and gender identity. Additionally, the exposure to women shifted men’s occupational choices towards jobs with a higher proportion of women. The results of this study show that men's gendered attitudes are not fixed and can change through exposure with women. However, the study also found that these short-term changes in men’s attitudes toward women did not persist in the long run. Men's gender attitudes during successive military service deteriorated, and there were differences to the control group. There was also no effect on men's education, occupation or workplace choices with a larger fraction of women in the long run. The researchers stated that the relatively short duration of the intervention during boot camp might explain these results.

The implications of this study suggest that exposure to women can be an effective strategy to overcome biases and stereotypical attitudes regarding women's suitability for jobs and gender roles in society. However, to maintain these improvements, continuous and intensive contact with women may be necessary. The study also claims that integrating women into traditionally male-dominated environments like the military can be achieved without compromising camaraderie, team effectiveness, or preparedness for future service. (contrary to existing beliefs of policy makers)

The article further states that there are still unanswered questions in this field which is whether the study’s results transfer to other settings.

2. Discussion:

The fundamental problems of causal inference in general and in regards to this paper specifically is the possibility of reverse causality, self-selection bias and unobserved heterogeneity. Concerning reverse causality, it is challenging to determine whether gender attitudes influence occupational segregation or if occupational segregation itself influences gender attitudes. For example, working in a male-dominated occupation could lead men to develop less egalitarian attitudes. However, it is also possible that men with pre-existing less egalitarian attitudes self-select into such occupations. Such self-selection bias can confound the causal relationship between gender attitudes and occupational segregation. Another issue is that there may be some unobserved heterogeneity meaning underlying variables that affect both gender attitudes and occupational segregation but are not directly observed or measured through variables. These unobserved variables can further introduce bias and complicate the identification of a causal effect.

The author's strategy to estimate causality involves using a randomized field experiment in the Norwegian military context to estimate the effects of integrating female recruits into all-male teams. By collecting data through surveys at different time points (three survey waves in total, on pre-existing gender attitudes, before and after the boot camp) and comparing the treated group with the control group the authors claim to estimate the causal effect of the impact of gender integration on gender attitudes.

The authors' field experiment implies various typical assumptions to identify causal inference.

First, the experiment assumes that the exposure to female recruits in mixed-gender squads does not have any indirect effects on gender attitudes in all-male squads. Second, the experiment assumes no spill-over effects between the treated group and the control group. Furthermore, it assumes that the recruits who participated in the survey are representative regarding gender attitudes toward women in the overall population of recruits. (by assuming there is no selection bias in the baseline survey) Additionally, it assumes that the random assignment of recruits to rooms persists throughout the initialisation boot camp – even though recruits may switch their room during the boot camp this is rather rarely the case according to Norwegian military officials. Last but not at least, the field experiment assumes that the effects observed in men's attitudes toward gender integration do persist over a six-month period into military service. Other assumptions are honesty in survey responses by acknowledging the possibility of a social desirability bias and reducing its influence by ensuring anonymity, emphasizing academic research purposes and assumptions of causal

inference (that the random assignment of female recruits to squads during boot camp allows for identifying a causal effect) and general comparability between the treatment and control group. Despite this sophisticated setting of the conducted experiment, field experiments are in general considered to have poor reliability. First of all, despite random assignment to squads, there could still be other unobserved confounders that affect the attitudes of the participants. (implies possibility of reverse causality) Secondly, the specific conditions and context of the field experiment are difficult to reproduce and therefore are limiting the generalizability of the author's findings. Furthermore, the field experiment described does not involve random allocation of participants to experimental conditions. Contrary, the squads were pre-set, and the randomization occurred at the troop level. This can lead towards potential sample bias, as the composition of squads may already be influenced by unobserved variables that in the experiment were not accounted for.

Regarding the experiment's setting it may seem extremely difficult to identify a causal effect since the author's also claim that there are contrary research findings, studies and contrary existing beliefs of policy makers regarding their main research question.

To summarize the authors, use a (partially) randomized field experiment design in which female recruits are randomly assigned to some squads while excluded from others during the eight-week boot camp to. A random assignment at the troop level is used to measure and estimate the causal effect of exposure to female recruits towards men's gender attitudes. By comparing the treatment group (men in squads with female recruits) with the control group (men in all-male squads), the authors claim to measure the causal effect. Data is collected through surveys conducted before and after the treatment (survey on gender attitudes, baseline survey and survey at the end of boot camp). The regression model is specified as $y_{ij2} = \alpha + \beta y_{i1} + \theta F_{ij1} + \gamma x_{i1} + \epsilon_{ij2}$ and includes a regression term for fixed effects for troops, controls for baseline attitudes (y_{i1}), a dummy variable for the presence of a female recruits in the squad (F_{ij1}), and control variables (x_{i1}). The inclusion troop fixed effects, the model controls for unobserved heterogeneity at the troop level and the inclusion of baseline attitudes as controls serves for individual-specific attitudes before the treatment. Furthermore, the author's address concerns about experimenter demand effects and social desirability bias. The recruits were not informed that the random assignment of female recruits served as part of an experiment. The surveys were conducted in a large group setting, not at the squad level, reducing social desirability bias. Respectively, the surveys included gender questions and many unrelated questions, further reducing potential bias.

The experiment design, regression modelling, treatment effect estimation, robustness checks, and considerations of potential biases should contribute to the authors' ability to establish a causal effect.

The authors include robustness checks by including additional covariates (captured by $xi1$) in the regression model to show that the treatment effects are consistent and unaffected. This To conclude, the authors provide a sophisticated field experiment with a randomized assignment of female recruits to squads (at the troop level) to measure the treatment effect. The authors' regression model includes fixed effects, baseline attitudes as controls, and additional covariates. The treatment effects reported in the paper show changes in men's gender attitudes due to exposure to women during boot camp as statistically significant. These findings are then further supported by multiple measures and inclusion of robustness checks.

While the paper presents a overall convincing analysis, there may still be potential limitations or unaddressed concerns that could impact the interpretation the author's finding.

Some possible considerations include:

- i. **Generalizability:** The study focuses on the military in Norway, where the integration of women had already been appearing for several years. The authors' findings may not fully identify or measure the effects of initial integration or the experiences of other contexts.
- ii. **Long-run effects:** The study examines short-term effects immediately after boot camp and six months into subsequent military service. The long-term effect of the observed attitude changes and their impact on behaviour or outcomes after the military service period remain unknown.
- iii. **Measurement errors:** The paper uses survey responses to measure gender attitudes. Survey data may introduce social desirability bias and may not always align with real-world situations.
- iv. **External validity:** This paper focuses on the effects of exposure to women during the initialisation boot camp. The applicability of the findings to other contexts, such as other may require further investigation also in relation to the sample selection bias.
- v. **Unobserved confounders:** Despite the regression model controlling for baseline attitudes ($yi1$) and additional variables ($xi1$), there may still be unobservable variables that could influence both squad assignment and attitudes.

Concerning the appropriateness of the model, regression models and OLS estimation are commonly used methods in the field of econometrics for identifying causal relationships.

Regarding the survey data, the author's claim that they were not directly involved in conducting the surveys which is in general considered as good practice for conduction surveys.

Overall the results presented in the paper that suggest positive (contrary to policy makers' existing beliefs as claimed by the authors) effects of the integration of women into military squads on male recruits' performance, satisfaction, camaraderie, team effectiveness, or preparedness. Nevertheless, it is difficult to generalize the authors' findings or using them to make policy recommendations for other military or contexts.

3. Comparison:

Firstly, the authors cited surveys conducted by the Norwegian Armed Forces and the Norwegian Defense Research Establishment in 2008 which focused on recruitment, motivation, and attitudes toward gender integration among soldiers. The results of these surveys proposed that integrating female soldiers into the military was largely successful – even though their analysis was based on non-random assignment of women to squads. The authors also cited anthropological studies Hellum et al. (2014, 2017) and Ellingsen et al. (2014) in which the effects of mixed-gender rooms on gender attitudes and behaviours were investigated. The first study concluded that mixed-gender rooms increased feelings of sameness between genders and the second study found that mixed rooms increased understanding, desexualisation, and reduced the risk of sexual harassment. Additionally, the authors also cited a qualitative study by Ellingsen et al. (2016) which findings were not discussed, but used to highlight the further research that supports the positive effects of gender integration in the military.

The mentioned studies support the positive effects of gender integration and the authors' main findings regarding the absence of negative effects on camaraderie, team effectiveness, and preparedness in the Norwegian military.

4. Conclusion:

Reading the author's paper served as useful refresher of methods used in econometrics and causal inference by covering interpretation of research designs, regression models, treatment effects, survey data, and data analysis skills. (e.g. interpretation of descriptive statistics, regression coefficients, charts, etc.) In terms of this author's research question, this task has enhanced my academic knowledge and increased awareness of the effect of gender

integration by exposure on gender attitudes, even though the authors' findings of this study may not apply in other contexts than the Norwegian military.

5. Sources:

- Gordon B Dahl, Andreas Kotsadam, and Dan-Olof Rooth. "Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams". In: *The Quarterly Journal of Economics* 136.2 (2021), pp. 987–1030. DOI: 10.1093/qje/qjaa047