

A. Supplementary material

A.1. Statistical significance of results

Each bias detection experiment was performed three times, due to the inherent randomness of the permutation test. We report the differences of the computed p -values (note that the d -values are exact) between the three runs: in 93% of cases the differences are in the range of ± 0.005 , with the maximum difference of ± 0.0135 observed in the first layer of the supervised model (Intersectional-Valence-BFBM).

A.2. Bias detections in the random model

Complementing Section 4.2 of the full paper, we provide further examples of bias detections in the randomly initialized ResNet models. At least 10 biases are detected in most ResNet layers of the 3 randomly initialized models (see Tables 1 and 7), namely:

- Weapon - Tool vs. Black - White
- Weapon - Tool (modern) vs. Black - White
- Thin - Overweight vs. Pleasant - Unpleasant
- Light skin tone - Dark skin tone vs. Pleasant - Unpleasant
- Lincoln - Trump vs. Pleasant - Unpleasant
- Arab (Muslim) - Other people vs. Pleasant - Unpleasant
- Career - Family vs. Male(White) - Female(White)
- Career - Family vs. Male(White+Black) - Female(White+Black)
- Career - Family vs. Male(Black) - Female(Black)
- Career - Family vs. Male(Black) - Female(White)

We hypothesize that the source of these biases is the distribution of the pixel values rather than the concepts represented in the images. To exemplify this hypothesis we refer to one of the biases detected by the random model in the GAP-layer: Lincoln - Trump vs. Pleasant - Unpleasant. As can be seen in Tables 1 and 7, this bias is present in the GAP-layer embeddings of the random model, with $p = 0.05$. After randomly shuffling the pixels in the images representing the two presidents, as shown on Figures 1 and 2, we repeat the bias detection process and observe that this bias is still present, with $p = 0.004$. Therefore, not only the bias persisted but also it became much more statistically significant. We can argue that this bias detection is triggered by the correlation between the pixel values distributions of the test data and not by the semantic concepts represented by the data.

Table 1. Comparison between the results of the bias detection in the GAP embeddings of the random model, performed on the transformed images and the original images (see Section A.2 of this supplementary material). Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Transformed images	Original images
Weight	<0.01 1.8	<0.01 1.88
Skin-Tone	0.03 1.01	<0.01 1.4
President - Bush vs. Trump	0.06 0.91	0.44 0.08
President - Lincoln vs. Trump	<0.01 1.4	0.05 0.98
Gender-Career	0.76 -0.16	<0.01 0.62
Intersectional-Gender-Science-WMBM	<0.01 0.92	0.01 0.84
Intersectional-Gender-Science-WMBF	<0.01 1.03	<0.01 1.02
Intersectional-Gender-Career-MF	0.87 -0.25	0.08 0.31
Intersectional-Gender-Career-WMBM	0.19 0.28	0.06 0.48
Intersectional-Gender-Career-WMWF	0.99 -0.81	0.06 0.5
Intersectional-Gender-Career-BMWF	0.99 -0.96	<0.01 0.89
Intersectional-Valence-BW	<0.01 1.29	<0.01 0.95
Intersectional-Valence-WMBM	<0.01 0.85	0.04 0.54
Intersectional-Valence-WMBF	<0.01 1.1	0.01 0.75
Intersectional-Valence-WMWF	0.01 0.77	0.02 0.63
Intersectional-Valence-WFBM	<0.01 1.49	<0.01 1.21
Intersectional-Valence-WFBF	<0.01 1.59	<0.01 1.36
Intersectional-Valence-FM	0.09 0.29	0.23 0.17

Table 1 comprises the differences in bias detection after pixel shuffling (transformed images) for all the detected biases in the GAP-layer embeddings.

We observe that for the majority of biases detected in the embeddings of the original images, biases are still detected when the images are transformed.

A.3. Bias detections in shallow layers

Overall, 17 biases (at $p_t \leq 0.05$) are detected for at least 8 out of 12 considered models (excluding the random model) in the first ResNet block (See Table 2). The experiments presented in Section A.2 of this supplementary material show

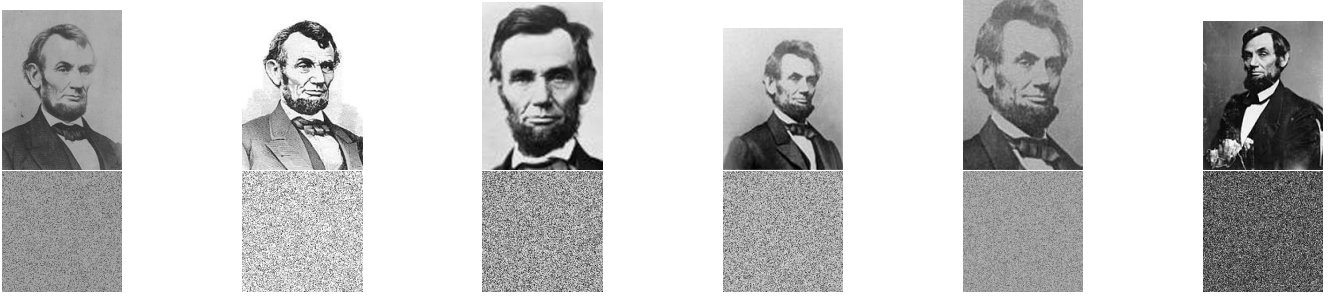


Figure 1. Images representing President Lincoln (on the top) and their transformed versions with all pixels randomly permuted (on the bottom).

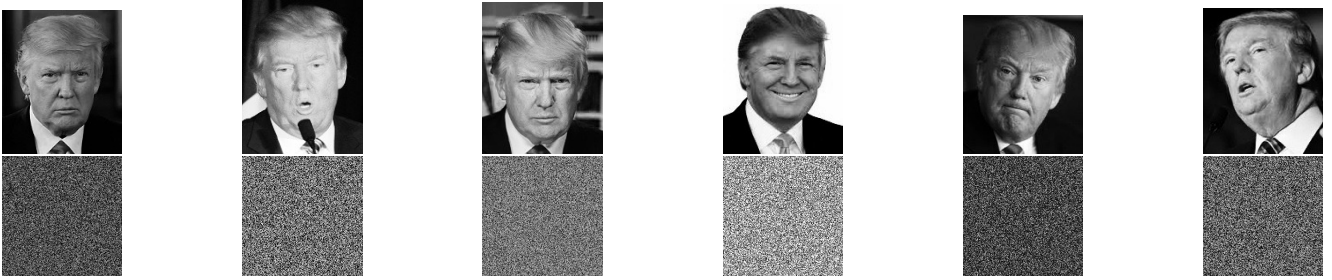


Figure 2. Images representing President Trump (on the top) and their transformed versions with all pixels randomly permuted (on the bottom).

that biases might be detected based on low-level similarities between corresponding images. Hence, we argue that a reason for the large number of bias detections in the shallow layers of ResNet relates to the correlations between the low-level features of the test images and not to the high-level abstract concepts that they represent.

A.4. Nature of the acquired biases

As was highlighted in the full paper, the general trend for all models is to incorporate more biases, as the depth of the embeddings (and, therefore, their level of semantic abstraction) increases. The trend, however, does not include the embeddings from the shallow layers, as they yield a large amount of biases due to the correlations in the low-level image characteristics (as described in Section A.3 of this supplementary material).

Figure 9 illustrates the biases detected in the embeddings extracted from different layers of one of the studied models, BYOL (for other models, see corresponding figures in Section A.5 of this supplementary material). One can observe that biases such as "Insect-Flower", "Intersectional-Gender-Career-WMBM", "Gender-Science" appear in the deep embeddings, while being absent in the shallow ones. Moreover, most of the detected biases (in all layers) correspond to the Intersectional biases (a similar trend is seen in other models).

A.5. Additional figures and tables

This Section contains figures and tables presenting the complete results of the bias detection experiments, specifically:

- Full results (p - and d -values) for the bias detection experiments (Tables 2, 3, 4, 5, 6, 7).
- Statistical strength of detected bias on a per-layer basis (Figures 3, 4, 5, 6, 7, 8).
- Presence of biases in different layers of the same model (Figures 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21).

Apart from the general conclusions and observations outlined in the main paper, one can see from Tables 3-7 that certain biases are present in different layers of all (or almost all) analyzed models. In particular, the "Race", "Intersectional-Valence-BW" and "Intersectional-Valence-WFBM" biases are present in the second block embeddings of all models (except the random model). Furthermore, "Intersectional-Valence-BW", "Intersectional-Valence-WFBM" and "Intersectional-Valence-WFBF" biases are found in the embeddings of the third block of all models (except the random model). A similar trend is observed in the deeper embeddings, which might signal that the training dataset (ImageNet) implicitly incorporates the aforementioned biases. A research in this direction (investigating not

yet recognized ImageNet biases and studying presence of biases across different models) opens paths for the future work.

Table 2. Results for the detection of biases based on the 1st block embeddings of 13 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.99 -0.96	0.99 -0.92	0.99 -0.86	0.99 -0.94	0.99 -0.87	0.99 -0.85	0.99 -0.65	0.99 -0.88	0.99 -0.82	0.99 -0.75	0.99 -0.65	0.99 -0.83	0.99 -0.8
Weapon	<0.01 1.62	0.94 -0.89	0.97 -1.02	0.96 -0.96	0.95 -0.92	0.97 -1.07	0.99 -1.28	0.96 -0.96	0.97 -1.02	0.98 -1.18	0.99 -1.27	0.97 -1.02	0.97 -1.06
Weapon (Modern)	0.02 1.13	0.96 -0.99	0.95 -0.98	0.91 -0.78	0.97 -1.06	0.92 -0.8	0.8 -0.51	0.97 -1.05	0.95 -0.97	0.76 -0.43	0.75 -0.4	0.95 -0.96	0.92 -0.81
Native	0.65 -0.2	0.77 -0.38	0.7 -0.27	0.58 -0.11	0.8 -0.43	0.43 0.09	0.26 0.33	0.74 -0.34	0.63 -0.17	0.38 0.16	0.39 0.14	0.57 -0.09	0.53 -0.03
Asian	0.85 -0.62	0.93 -0.85	0.93 -0.86	0.94 -0.88	0.93 -0.85	0.94 -0.93	0.97 -1.04	0.93 -0.88	0.95 -0.95	0.95 -0.98	0.97 -1.09	0.94 -0.92	0.95 -0.94
Weight	<0.01 1.73	<0.01 1.89	<0.01 1.85	<0.01 1.84	<0.01 1.89	<0.01 1.83	<0.01 1.37	<0.01 1.89	<0.01 1.85	<0.01 1.66	<0.01 1.39	<0.01 1.85	<0.01 1.82
Skin-Tone	0.01 1.1	<0.01 1.64	<0.01 1.59	<0.01 1.51	<0.01 1.67	<0.01 1.54	0.06 0.83	<0.01 1.65	<0.01 1.57	<0.01 1.29	0.07 0.82	<0.01 1.58	<0.01 1.51
Disability	0.31 0.37	0.26 0.56	0.19 0.64	0.19 0.7	0.27 0.61	0.37 0.27	0.29 0.45	0.17 0.7	0.27 0.54	0.29 0.47	0.24 0.47	0.2 0.66	0.21 0.64
President - Kennedy vs. Trump	0.47 0.04	0.44 0.1	0.63 -0.2	0.66 -0.25	0.41 0.14	0.51 -0.01	0.85 -0.67	0.46 0.06	0.6 -0.14	0.71 -0.37	0.85 -0.66	0.55 -0.08	0.69 -0.3
President - B. Clinton vs. Trump	0.19 0.55	0.65 -0.23	0.68 -0.3	0.68 -0.28	0.65 -0.24	0.49 0.01	0.41 0.15	0.64 -0.21	0.55 -0.09	0.5 -0.01	0.37 0.23	0.64 -0.22	0.64 -0.23
President - Bush vs. Trump	0.52 -0.03	0.25 0.41	0.27 0.37	0.29 0.33	0.24 0.43	0.26 0.39	0.37 0.23	0.24 0.42	0.27 0.37	0.3 0.3	0.37 0.24	0.27 0.36	0.3 0.33
President - Lincoln vs. Trump	0.01 1.19	0.02 1.25	0.02 1.25	0.02 1.24	0.02 1.25	0.01 1.31	0.02 1.09	0.01 1.28	0.02 1.29	0.01 1.24	0.02 1.07	0.02 1.28	0.02 1.25
Religion	0.86 -0.64	0.59 -0.13	0.53 -0.05	0.56 -0.09	0.57 -0.11	0.51 -0.02	0.49 0.01	0.55 -0.08	0.53 -0.04	0.51 -0.01	0.48 0.03	0.56 -0.09	0.53 -0.04
Sexuality	0.58 -0.1	0.5 -0.01	0.46 0.06	0.45 0.06	0.5 -0.0	0.41 0.12	0.22 0.38	0.46 0.04	0.41 0.11	0.28 0.28	0.24 0.34	0.4 0.13	0.38 0.14
Race	0.1 0.75	0.84 -0.62	0.74 -0.4	0.69 -0.32	0.86 -0.68	0.69 -0.3	0.1 0.76	0.84 -0.62	0.72 -0.37	0.4 0.13	0.11 0.72	0.74 -0.4	0.63 -0.2
Arab-Muslim	0.01 0.95	0.12 0.53	0.12 0.53	0.12 0.54	0.13 0.51	0.12 0.51	0.16 0.45	0.12 0.53	0.13 0.51	0.14 0.5	0.18 0.42	0.12 0.53	0.14 0.52
Age	0.56 -0.09	0.56 -0.09	0.52 -0.03	0.49 0.01	0.58 -0.13	0.49 0.01	0.28 0.36	0.57 -0.1	0.5 -0.0	0.35 0.24	0.24 0.43	0.51 -0.01	0.46 0.07
Gender-Science	0.14 0.26	0.83 -0.21	0.71 -0.13	0.85 -0.24	0.75 -0.16	0.81 -0.2	0.8 -0.19	0.78 -0.17	0.82 -0.2	0.76 -0.16	0.79 -0.18	0.81 -0.2	0.65 -0.09
Gender-Career	<0.01 0.62	0.01 0.51	<0.01 0.63	0.02 0.44	<0.01 0.56	0.01 0.54	0.05 0.36	0.01 0.57	0.01 0.52	0.03 0.44	0.02 0.45	0.01 0.5	<0.01 0.59
Intersectional-Gender-Science-MF	0.71 -0.13	0.98 -0.45	0.98 -0.46	0.98 -0.46	0.97 -0.45	0.99 -0.54	0.99 -0.58	0.98 -0.47	0.99 -0.51	0.99 -0.54	0.99 -0.57	0.99 -0.5	0.98 -0.46
Intersectional-Gender-Science-WMBM	0.31 0.15	<0.01 0.92	<0.01 0.83	<0.01 0.81	<0.01 0.94	0.01 0.78	0.04 0.55	<0.01 0.91	<0.01 0.82	0.02 0.64	0.05 0.51	<0.01 0.85	0.01 0.76
Intersectional-Gender-Science-WMBF	0.15 0.34	<0.01 0.85	<0.01 0.77	<0.01 0.76	<0.01 0.86	0.02 0.66	0.09 0.44	<0.01 0.82	0.01 0.72	0.04 0.54	0.1 0.4	0.01 0.75	0.01 0.72
Intersectional-Gender-Science-WMWF	0.92 -0.46	0.99 -0.98	0.99 -0.95	0.99 -0.98	0.99 -0.97	0.99 -1.04	0.99 -1.03	0.99 -0.99	0.99 -1.03	0.99 -1.0	0.99 -1.02	0.99 -1.01	0.99 -0.93
Intersectional-Gender-Science-BMBF	0.33 0.14	0.62 -0.1	0.61 -0.08	0.57 -0.06	0.63 -0.11	0.7 -0.17	0.68 -0.15	0.63 -0.11	0.66 -0.13	0.66 -0.14	0.67 -0.14	0.66 -0.13	0.58 -0.06
Intersectional-Gender-Science-BMWF	0.96 -0.58	0.99 -1.59	0.99 -1.56	0.99 -1.57	0.99 -1.6	0.99 -1.59	0.99 -1.56	0.99 -1.6	0.99 -1.6	0.99 -1.55	0.99 -1.54	0.99 -1.6	0.99 -1.54
Intersectional-Gender-Career-MF	0.03 0.42	0.01 0.55	<0.01 0.55	0.05 0.36	<0.01 0.61	<0.01 0.59	0.03 0.41	<0.01 0.58	<0.01 0.58	0.02 0.46	0.03 0.43	<0.01 0.61	0.01 0.52
Intersectional-Gender-Career-WMBM	0.38 0.09	0.38 0.1	0.35 0.11	0.66 -0.14	0.19 0.27	0.4 0.09	0.45 0.04	0.28 0.18	0.31 0.16	0.36 0.11	0.43 0.06	0.18 0.3	0.38 0.1
Intersectional-Gender-Career-WMBF	0.09 0.44	0.04 0.54	0.05 0.51	0.01 0.75	0.14 0.34	0.02 0.68	0.04 0.56	0.05 0.52	0.04 0.55	0.06 0.49	0.05 0.53	0.16 0.32	0.05 0.52
Intersectional-Gender-Career-WMWF	0.12 0.39	0.02 0.66	0.01 0.69	0.22 0.25	<0.01 0.81	0.01 0.71	0.17 0.31	0.01 0.76	0.01 0.7	0.09 0.44	0.12 0.38	<0.01 0.81	0.03 0.57
Intersectional-Gender-Career-BMBF	0.08 0.46	0.06 0.5	0.06 0.49	0.08 0.45	0.05 0.51	0.04 0.56	0.06 0.49	0.04 0.54	0.04 0.54	0.06 0.5	0.07 0.48	0.05 0.51	0.06 0.5
Intersectional-Gender-Career-BMWF	0.11 0.4	0.03 0.56	0.03 0.59	0.49 0.01	<0.01 0.84	0.05 0.53	0.2 0.27	0.01 0.68	0.03 0.6	0.09 0.44	0.14 0.34	<0.01 0.84	0.05 0.51
Intersectional-Valence-BW	0.2 0.2	<0.01 1.28	<0.01 1.28	<0.01 1.25	<0.01 1.3	<0.01 1.24	<0.01 1.14	<0.01 1.29	<0.01 1.26	<0.01 1.21	<0.01 1.13	<0.01 1.28	<0.01 1.27
Intersectional-Valence-WMBM	0.43 0.06	<0.01 0.98	<0.01 0.99	<0.01 0.94	<0.01 1.01	<0.01 0.95	<0.01 0.84	<0.01 1.0	<0.01 0.96	<0.01 0.92	<0.01 0.83	<0.01 1.0	<0.01 0.98
Intersectional-Valence-WMBF	0.3 0.18	<0.01 0.98	<0.01 0.97	<0.01 0.95	<0.01 1.0	<0.01 0.91	<0.01 0.82	<0.01 0.99	<0.01 0.95	<0.01 0.9	<0.01 0.8	<0.01 0.98	<0.01 0.95
Intersectional-Valence-WMWF	0.36 0.15	<0.01 0.91	<0.01 0.89	<0.01 0.88	<0.01 0.91	<0.01 0.92	<0.01 0.79	<0.01 0.91	<0.01 0.91	<0.01 0.83	<0.01 0.82	<0.01 0.89	<0.01 0.89
Intersectional-Valence-WFBM	0.24 0.24	<0.01 1.59	<0.01 1.58	<0.01 1.56	<0.01 1.59	<0.01 1.57	<0.01 1.46	<0.01 1.59	<0.01 1.58	<0.01 1.53	<0.01 1.47	<0.01 1.59	<0.01 1.57
Intersectional-Valence-BFBM	0.66 -0.13	0.47 0.02	0.49 0.0	0.51 -0.01	0.48 0.01	0.41 0.06	0.51 -0.01	0.49 0.01	0.51 0.01	0.48 0.02	0.49 0.01	0.46 0.03	0.47 0.02
Intersectional-Valence-WFBF	0.14 0.38	<0.01 1.59	<0.01 1.56	<0.01 1.56	<0.01 1.59	<0.01 1.55	<0.01 1.42	<0.01 1.58	<0.01 1.56	<0.01 1.51	<0.01 1.42	<0.01 1.57	<0.01 1.55
Intersectional-Valence-FM	0.5 -0.0	0.05 0.37	0.05 0.36	0.05 0.35	0.05 0.37	0.03 0.4	0.07 0.33	0.05 0.37	0.05 0.37	0.06 0.35	0.05 0.35	0.05 0.37	0.05 0.37

Table 3. Results for the detection of biases based on the 2nd block embeddings of 13 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.99 -0.74	0.98 -0.47	0.99 -0.59	0.97 -0.47	0.73 -0.14	0.99 -0.53	0.98 -0.47	0.98 -0.48	0.99 -0.57	0.99 -0.57	0.97 -0.45	0.99 -0.64	0.97 -0.47
Weapon	<0.01 1.6	0.99 -1.45	0.84 -0.61	0.99 -1.26	0.78 -0.49	0.58 -0.12	0.87 -0.69	0.99 -1.14	0.66 -0.27	0.93 -0.82	0.73 -0.39	0.82 -0.58	0.97 -1.06
Weapon (Modern)	0.03 1.11	0.99 -1.72	0.99 -1.3	0.99 -1.6	0.99 -1.59	0.78 -0.46	0.91 -0.81	0.79 -0.49	0.51 -0.02	0.92 -0.83	0.79 -0.47	0.86 -0.64	0.95 -0.97
Native	0.74 -0.34	0.73 -0.32	0.48 0.02	0.94 -0.78	0.8 -0.42	0.83 -0.49	0.59 -0.12	0.71 -0.29	0.78 -0.41	0.84 -0.52	0.84 -0.51	0.8 -0.45	0.87 -0.58
Asian	0.85 -0.63	0.64 -0.24	0.58 -0.13	0.69 -0.31	0.71 -0.36	0.61 -0.19	0.41 0.15	0.88 0.7	0.67 -0.27	0.49 0.02	0.48 0.04	0.36 0.21	0.63 -0.21
Weight	<0.01 1.76	<0.01 1.17	0.16 0.47	<0.01 1.12	0.1 0.59	0.03 0.84	0.04 0.79	0.02 0.89	0.01 1.04	0.07 0.67	0.02 0.86	0.06 0.72	0.07 0.66
Skin-Tone	0.03 0.98	0.37 0.18	0.69 -0.29	0.26 0.36	0.62 -0.19	0.4 0.14	0.55 -0.07	0.49 0.01	0.51 -0.02	0.68 -0.27	0.61 -0.15	0.65 -0.23	0.53 -0.04
Disability	0.27 0.5	0.04 1.23	0.16 0.77	0.09 1.12	0.1 1.05	0.09 0.93	0.1 1.09	0.04 1.22	0.09 1.08	0.1 0.96	0.11 1.06	0.11 1.04	0.07 1.13
President - Kennedy vs. Trump	0.43 0.11	0.95 -0.94	0.95 -0.95	0.95 -0.97	0.98 -1.15	0.92 -0.83	0.94 -0.92	0.96 -1.0	0.95 -0.93	0.92 -0.81	0.95 -0.93	0.93 -0.87	0.95 -0.94
President - B. Clinton vs. Trump	0.17 0.6	0.25 0.41	0.4 0.16	0.34 0.26	0.36 0.23	0.16 0.58	0.43 0.11	0.22 0.46	0.29 0.32	0.45 0.08	0.42 0.12	0.59 -0.12	0.39 0.18
President - Bush vs. Trump	0.43 0.1	0.32 0.31	0.43 0.11	0.41 0.14	0.4 0.15	0.33 0.28	0.35 0.23	0.34 0.26	0.39 0.17	0.38 0.19	0.34 0.21	0.43 0.11	0.35 0.23
President - Lincoln vs. Trump	0.01 1.22	0.11 0.76	0.18 0.54	0.13 0.7	0.28 0.35	0.09 0.8	0.27 0.4	0.13 0.67	0.12 0.7	0.14 0.63	0.21 0.47	0.16 0.58	0.17 0.55
Religion	0.85 -0.59	0.66 -0.23	0.48 0.02	0.7 -0.3	0.59 -0.12	0.66 -0.22	0.68 -0.26	0.69 -0.28	0.7 -0.3	0.61 -0.16	0.7 -0.3	0.68 -0.26	0.57 -0.1
Sexuality	0.48 0.02	0.33 0.22	0.23 0.36	0.32 0.23	0.09 0.63	0.21 0.38	0.2 0.41	0.29 0.28	0.26 0.31	0.18 0.43	0.19 0.43	0.25 0.33	0.3 0.27
Race	0.12 0.69	<0.01 1.59	<0.01 1.63	<0.01 1.59	<0.01 1.59	<0.01 1.69	<0.01 1.5	<0.01 1.61	<0.01 1.66	<0.01 1.69	<0.01 1.6	<0.01 1.63	<0.01 1.62
Arab-Muslim	0.01 0.97	0.14 0.5	0.12 0.53	0.17 0.45	0.1 0.59	0.16 0.45	0.06 0.7	0.18 0.43	0.18 0.41	0.07 0.66	0.07 0.65	0.05 0.71	0.12 0.54
Age	0.59 -0.15	0.18 0.6	0.21 0.54	0.23 0.45	0.18 0.61	0.11 0.74	0.29 0.42	0.12 0.72	0.16 0.63	0.15 0.65	0.18 0.6	0.22 0.51	0.17 0.6
Gender-Science	0.47 0.05	0.49 0.01	0.5 0.0	0.59 -0.05	0.31 0.11	0.45 0.03	0.25 0.15	0.57 -0.04	0.5 0.0	0.29 0.12	0.35 0.08	0.32 0.1	0.59 -0.05
Gender-Career	0.01 0.57	0.29 0.12	0.03 0.41	0.08 0.3	0.05 0.36	0.13 0.26	0.08 0.31	0.17 0.22	0.21 0.18	0.08 0.32	0.12 0.27	0.2 0.19	0.2 0.2
Intersectional-Gender-Science-MF	0.78 -0.18	0.99 -0.49	0.97 -0.42	0.99 -0.53	0.96 -0.39	0.99 -0.47	0.99 -0.52	0.99 -0.48	0.99 -0.52	0.99 -0.56	0.99 -0.53	0.99 -0.49	0.99 -0.53
Intersectional-Gender-Science-WMBM	0.2 0.27	0.17 0.31	0.46 0.04	0.21 0.27	0.4 0.09	0.32 0.16	0.12 0.38	0.17 0.31	0.2 0.28	0.21 0.27	0.23 0.25	0.18 0.3	0.2 0.28
Intersectional-Gender-Science-WMBF	0.1 0.4	0.24 0.23	0.4 0.09	0.32 0.16	0.34 0.15	0.29 0.19	0.25 0.23	0.22 0.26	0.26 0.21	0.36 0.12	0.35 0.14	0.26 0.22	0.35 0.13
Intersectional-Gender-Science-WMWF	0.96 -0.57	0.99 -0.82	0.99 -0.77	0.99 -0.84	0.99 -0.74	0.99 -0.9	0.99 -0.83	0.99 -0.85	0.99 -0.88	0.99 -0.86	0.99 -0.86	0.99 -0.81	0.99 -0.82
Intersectional-Gender-Science-BMBF	0.39 0.09	0.63 -0.12	0.41 0.08	0.7 -0.17	0.41 0.07	0.45 0.04	0.73 -0.21	0.58 -0.06	0.62 -0.1	0.74 -0.22	0.68 -0.15	0.64 -0.12	0.74 -0.21
Intersectional-Gender-Science-BMWF	0.99 -0.76	0.99 -1.26	0.99 -1.05	0.99 -1.31	0.99 -1.05	0.99 -1.27	0.99 -1.32	0.99 -1.29	0.99 -1.34	0.99 -1.33	0.99 -1.31	0.99 -1.29	0.99 -1.24
Intersectional-Gender-Career-MF	0.04 0.41	0.77 -0.17	0.31 0.12	0.7 -0.12	0.42 0.04	0.74 -0.14	0.82 -0.21	0.71 -0.12	0.72 -0.13	0.8 -0.19	0.81 -0.2	0.75 -0.15	0.73 -0.14
Intersectional-Gender-Career-WMBM	0.39 0.09	0.83 -0.3	0.44 0.05	0.83 -0.32	0.57 -0.06	0.86 -0.35	0.89 -0.4	0.84 -0.32	0.86 -0.35	0.73 -0.2	0.82 -0.3	0.79 -0.27	0.73 -0.2
Intersectional-Gender-Career-WMBF	0.1 0.42	0.09 0.43	0.32 0.15	0.07 0.45	0.18 0.3	0.08 0.46	0.06 0.49	0.06 0.48	0.06 0.48	0.25 0.22	0.12 0.38	0.11 0.4	0.21 0.27
Intersectional-Gender-Career-WMWF	0.14 0.36	0.92 -0.45	0.46 0.03	0.85 -0.34	0.67 -0.17	0.88 -0.39	0.93 -0.48	0.87 -0.39	0.88 -0.39	0.88 -0.37	0.93 -0.47	0.9 -0.42	0.84 -0.34
Intersectional-Gender-Career-BMBF	0.08 0.44	0.38 0.11	0.27 0.2	0.35 0.12	0.21 0.26	0.37 0.1	0.42 0.07	0.31 0.16	0.36 0.12	0.5 -0.0	0.41 0.08	0.34 0.14	0.42 0.06
Intersectional-Gender-Career-BMWF	0.12 0.39	0.99 -0.81	0.4 0.09	0.99 -0.76	0.79 -0.26	0.99 -0.83	0.99 -0.93	0.99 -0.82	0.99 -0.81	0.98 -0.65	0.99 -0.86	0.99 -0.76	0.98 -0.61
Intersectional-Valence-BW	0.31 0.12	<0.01 0.84	<0.01 0.78	<0.01 0.8	<0.01 0.88	<0.01 0.82	<0.01 0.94	<0.01 0.8	<0.01 0.79	<0.01 0.84	<0.01 0.9	<0.01 0.87	<0.01 0.78
Intersectional-Valence-WMBM	0.5 -0.01	0.06 0.49	0.09 0.44	0.08 0.45	0.04 0.56	0.07 0.46	0.02 0.65	0.09 0.44	0.09 0.42	0.05 0.52	0.04 0.57	0.04 0.58	0.08 0.45
Intersectional-Valence-WMBF	0.41 0.08	0.05 0.51	0.03 0.58	0.06 0.49	0.02 0.66	0.05 0.52	0.02 0.61	0.06 0.49	0.07 0.46	0.05 0.52	0.04 0.56	0.04 0.57	0.09 0.44
Intersectional-Valence-WMWF	0.35 0.17	0.02 0.63	0.05 0.52	0.03 0.58	0.04 0.56	0.01 0.67	0.02 0.65	0.02 0.66	0.01 0.68	0.02 0.63	0.01 0.67	0.03 0.6	0.02 0.62
Intersectional-Valence-WFBM	0.28 0.19	<0.01 1.21	<0.01 1.04	<0.01 1.18	<0.01 1.13	<0.01 1.17	<0.01 1.26	<0.01 1.15	<0.01 1.17	<0.01 1.18	<0.01 1.25	<0.01 1.21	<0.01 1.16
Intersectional-Valence-BFBM	0.61 -0.09	0.48 0.01	0.72 -0.19	0.56 -0.05	0.64 -0.12	0.57 -0.06	0.41 0.08	0.56 -0.05	0.57 -0.06	0.51 -0.0	0.45 0.04	0.49 0.01	0.47 0.02
Intersectional-Valence-WFBF	0.19 0.29	<0.01 1.28	<0.01 1.14	<0.01 1.22	<0.01 1.23	<0.01 1.22	<0.01 1.26	<0.01 1.21	<0.01 1.2	<0.01 1.17	<0.01 1.27	<0.01 1.19	<0.01 1.18
Intersectional-Valence-FM	0.44 0.03	0.08 0.32	0.24 0.16	0.11 0.28	0.19 0.2	0.09 0.29	0.06 0.34	0.1 0.3	0.09 0.31	0.09 0.3	0.06 0.34	0.1 0.29	0.08 0.31

Table 4. Results for the detection of biases based on the 3rd block embeddings of 13 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.98 -0.53	0.99 -0.72	0.99 -0.81	0.99 -0.66	0.99 -0.64	0.99 -0.68	0.99 -0.69	0.99 -0.69	0.99 -0.85	0.99 -0.89	0.99 -0.6	0.99 -0.69	0.99 -0.61
Weapon	<0.01 1.59	0.75 -0.42	0.75 -0.4	0.92 -0.82	0.69 -0.28	0.8 -0.51	0.07 0.89	0.66 -0.25	0.27 0.39	0.19 0.53	0.45 0.07	0.46 0.05	0.41 0.14
Weapon (Modern)	0.02 1.13	0.96 -0.97	0.88 -0.71	0.85 -0.61	0.67 -0.27	0.49 0.01	0.59 -0.14	0.72 -0.4	0.86 -0.65	0.89 -0.74	0.87 -0.68	0.89 -0.73	0.63 -0.2
Native	0.69 -0.28	0.57 -0.09	0.35 0.2	0.77 -0.38	0.7 -0.26	0.77 -0.4	0.53 -0.04	0.64 -0.19	0.62 -0.17	0.63 -0.19	0.56 -0.07	0.66 -0.23	0.42 0.11
Asian	0.8 -0.53	0.4 0.15	0.24 0.43	0.3 0.31	0.33 0.27	0.53 -0.06	0.15 0.61	0.38 0.2	0.57 -0.11	0.34 0.28	0.19 0.54	0.09 0.78	0.22 0.48
Weight	<0.01 1.73	0.01 1.03	0.04 0.78	0.02 0.88	0.05 0.73	0.03 0.83	0.41 0.12	0.02 0.91	0.01 0.99	0.05 0.74	0.14 0.49	0.04 0.77	<0.01 1.15
Skin-Tone	0.05 0.9	0.33 0.24	0.49 0.03	0.36 0.2	0.47 0.04	0.43 0.1	0.41 0.13	0.56 -0.08	0.39 0.16	0.43 0.1	0.43 0.1	0.53 -0.04	0.55 -0.06
Disability	0.49 0.04	0.09 1.05	0.04 1.22	0.09 1.07	0.2 0.68	0.04 1.24	0.29 0.49	0.07 1.12	0.56 -0.15	0.46 0.05	0.11 0.91	0.14 0.89	0.06 1.1
President - Kennedy vs. Trump	0.43 0.12	0.91 -0.78	0.92 -0.81	0.91 -0.78	0.89 -0.74	0.97 -1.06	0.98 -1.16	0.95 -0.96	0.91 -0.81	0.96 -0.99	0.99 -1.21	0.98 -1.12	0.9 -0.75
President - B. Clinton vs. Trump	0.26 0.46	0.61 -0.18	0.58 -0.13	0.54 -0.05	0.9 -0.76	0.73 -0.36	0.74 -0.37	0.67 -0.26	0.97 -1.03	0.73 -0.96	0.88 -0.36	0.66 -0.68	0.66 -0.26
President - Bush vs. Trump	0.51 -0.01	0.49 0.02	0.41 0.13	0.45 0.07	0.57 -0.11	0.4 0.16	0.35 0.22	0.41 0.13	0.76 -0.42	0.64 -0.22	0.46 0.07	0.56 -0.08	0.45 0.07
President - Lincoln vs. Trump	0.04 0.96	0.12 0.71	0.21 0.5	0.12 0.71	0.17 0.56	0.25 0.42	0.36 0.22	0.22 0.47	0.02 1.13	0.23 0.46	0.24 0.41	0.47 0.04	0.14 0.64
Religion	0.85 -0.59	0.57 -0.1	0.49 0.01	0.62 -0.16	0.52 -0.02	0.5 -0.0	0.69 -0.27	0.62 -0.17	0.58 -0.11	0.44 0.09	0.64 -0.19	0.48 0.02	0.43 0.09
Sexuality	0.49 0.02	0.38 0.15	0.34 0.19	0.39 0.14	0.23 0.37	0.35 0.19	0.33 0.21	0.33 0.21	0.22 0.38	0.34 0.2	0.31 0.24	0.36 0.19	0.2 0.4
Race	0.2 0.49	0.01 1.39	<0.01 1.5	<0.01 1.6	<0.01 1.45	<0.01 1.52	<0.01 1.56	<0.01 1.69	0.06 0.9	0.03 1.06	0.01 1.32	<0.01 1.49	<0.01 1.51
Arab-Muslim	0.01 0.91	0.06 0.68	0.05 0.72	0.07 0.67	0.05 0.76	0.08 0.63	0.06 0.69	0.1 0.59	0.09 0.6	0.06 0.71	0.06 0.71	0.04 0.81	0.05 0.73
Age	0.5 0.0	0.3 0.37	0.32 0.33	0.27 0.39	0.45 0.11	0.25 0.47	0.19 0.54	0.38 0.24	0.86 -0.69	0.75 -0.46	0.24 0.47	0.46 0.08	0.31 0.37
Gender-Science	0.57 -0.04	0.35 0.09	0.34 0.09	0.36 0.08	0.17 0.21	0.1 0.3	0.08 0.32	0.3 0.13	0.26 0.14	0.06 0.34	0.14 0.24	0.06 0.35	0.06 0.34
Gender-Career	0.01 0.55	0.05 0.36	0.03 0.42	0.09 0.31	0.02 0.46	0.15 0.23	0.08 0.31	0.03 0.43	0.33 0.1	0.08 0.31	0.04 0.39	0.04 0.4	0.01 0.53
Intersectional-Gender-Science-MF	0.71 -0.14	0.99 -0.58	0.99 -0.57	0.99 -0.57	0.99 -0.52	0.99 -0.54	0.99 -0.52	0.99 -0.55	0.99 -0.59	0.99 -0.57	0.99 -0.49	0.99 -0.52	0.99 -0.48
Intersectional-Gender-Science-WMBM	0.54 -0.03	0.21 0.27	0.33 0.15	0.2 0.27	0.23 0.24	0.21 0.27	0.55 -0.04	0.13 0.36	0.26 0.22	0.31 0.16	0.35 0.13	0.23 0.24	0.33 0.15
Intersectional-Gender-Science-WMBF	0.62 -0.12	0.44 0.05	0.54 -0.04	0.42 0.06	0.37 0.11	0.35 0.13	0.7 -0.17	0.3 0.17	0.49 0.01	0.6 -0.08	0.44 0.05	0.46 0.03	0.49 -0.0
Intersectional-Gender-Science-WMWF	0.77 -0.28	0.99 -0.81	0.99 -0.84	0.99 -0.81	0.99 -0.82	0.99 -0.87	0.99 -0.78	0.99 -0.82	0.99 -0.83	0.99 -0.77	0.99 -0.8	0.99 -0.74	0.99 -0.71
Intersectional-Gender-Science-BMBF	0.57 -0.06	0.84 -0.32	0.78 -0.24	0.83 -0.31	0.7 -0.18	0.7 -0.18	0.71 -0.18	0.8 -0.27	0.83 -0.31	0.85 -0.33	0.62 -0.1	0.8 -0.27	0.74 -0.2
Intersectional-Gender-Science-BMWF	0.66 -0.16	0.99 -1.26	0.99 -1.17	0.99 -1.27	0.99 -1.25	0.99 -1.31	0.99 -0.98	0.99 -1.32	0.99 -1.32	0.99 -1.17	0.99 -1.16	0.99 -1.17	0.99 -1.07
Intersectional-Gender-Career-MF	0.03 0.43	0.7 -0.12	0.4 0.06	0.61 -0.07	0.42 0.04	0.6 -0.06	0.26 0.15	0.47 0.01	0.57 -0.04	0.3 0.12	0.23 0.16	0.36 0.08	0.34 0.09
Intersectional-Gender-Career-WMBM	0.27 0.21	0.86 -0.35	0.75 -0.21	0.66 -0.13	0.8 -0.26	0.82 -0.3	0.71 -0.18	0.83 -0.33	0.75 -0.22	0.73 -0.2	0.84 -0.31	0.83 -0.32	0.59 -0.08
Intersectional-Gender-Career-WMBF	0.18 0.3	0.12 0.38	0.09 0.42	0.17 0.31	0.09 0.43	0.08 0.44	0.02 0.64	0.09 0.44	0.13 0.37	0.08 0.45	0.01 0.77	0.02 0.63	0.15 0.34
Intersectional-Gender-Career-WMWF	0.1 0.42	0.77 -0.26	0.62 -0.11	0.81 -0.29	0.57 -0.06	0.8 -0.27	0.73 -0.2	0.61 -0.1	0.74 -0.22	0.52 -0.03	0.66 -0.13	0.68 -0.15	0.61 -0.1
Intersectional-Gender-Career-BMBF	0.08 0.45	0.49 0.0	0.26 0.21	0.26 0.2	0.34 0.14	0.32 0.15	0.08 0.45	0.36 0.13	0.37 0.11	0.22 0.25	0.08 0.43	0.18 0.29	0.2 0.28
Intersectional-Gender-Career-BMWF	0.04 0.56	0.98 -0.67	0.86 -0.35	0.94 -0.49	0.88 -0.38	0.98 -0.62	0.88 -0.38	0.94 -0.51	0.92 -0.45	0.79 -0.26	0.93 -0.47	0.94 -0.49	0.74 -0.21
Intersectional-Valence-BW	0.56 -0.04	<0.01 0.75	<0.01 0.67	<0.01 0.69	<0.01 0.83	<0.01 0.83	0.01 0.57	<0.01 0.84	<0.01 0.68	<0.01 0.69	<0.01 0.74	<0.01 0.78	<0.01 0.8
Intersectional-Valence-WMBM	0.63 -0.12	0.1 0.41	0.14 0.34	0.14 0.34	0.07 0.48	0.07 0.48	0.24 0.23	0.07 0.47	0.18 0.3	0.12 0.38	0.13 0.36	0.07 0.46	0.07 0.47
Intersectional-Valence-WMBF	0.59 -0.08	0.12 0.38	0.14 0.35	0.21 0.26	0.05 0.51	0.08 0.46	0.21 0.27	0.05 0.51	0.21 0.26	0.24 0.23	0.07 0.47	0.11 0.39	0.06 0.5
Intersectional-Valence-WMWF	0.36 0.16	0.02 0.65	0.03 0.6	0.01 0.7	0.01 0.66	0.01 0.71	0.03 0.6	0.01 0.67	0.01 0.73	0.01 0.69	0.02 0.63	0.01 0.68	0.03 0.62
Intersectional-Valence-WFBM	0.48 0.02	<0.01 1.15	<0.01 1.03	<0.01 1.17	<0.01 1.22	<0.01 1.22	<0.01 0.88	<0.01 1.2	<0.01 1.18	<0.01 1.21	<0.01 1.03	<0.01 1.18	<0.01 1.13
Intersectional-Valence-BFBM	0.54 -0.04	0.45 0.04	0.51 -0.01	0.37 0.11	0.59 -0.08	0.45 0.04	0.54 -0.03	0.56 -0.05	0.44 0.05	0.28 0.19	0.62 -0.11	0.4 0.09	0.54 -0.04
Intersectional-Valence-WFBF	0.42 0.06	<0.01 1.15	<0.01 1.03	<0.01 1.12	<0.01 1.2	<0.01 1.23	<0.01 0.94	<0.01 1.23	<0.01 1.15	<0.01 1.05	<0.01 1.15	<0.01 1.13	<0.01 1.15
Intersectional-Valence-FM	0.41 0.06	0.05 0.35	0.09 0.31	0.03 0.42	0.09 0.3	0.04 0.38	0.09 0.31	0.07 0.32	0.03 0.41	0.02 0.45	0.1 0.28	0.04 0.39	0.1 0.29

Table 5. Results for the detection of biases based on the 4th block embeddings of 10 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.98 -0.52	0.99 -0.71	0.97 -0.46	0.98 -0.48	0.99 -0.76	0.99 -0.82	0.76 -0.17	0.99 -0.87	0.99 -0.99	0.99 -0.99	0.95 -0.4	0.91 -0.32	0.99 -0.84
Weapon	<0.01 1.52	0.13 0.67	0.22 0.46	0.16 0.62	0.45 0.05	0.62 -0.2	0.74 -0.38	0.75 -0.41	0.68 -0.29	0.83 -0.58	0.22 0.47	0.56 -0.09	0.69 -0.31
Weapon (Modern)	0.05 0.99	0.1 0.78	0.28 0.36	0.77 -0.46	0.72 -0.34	0.85 -0.62	0.09 0.78	0.83 -0.57	0.42 0.14	0.98 -0.19	0.3 0.33	0.27 0.39	0.96 -1.0
Native	0.82 -0.49	0.73 -0.33	0.58 -0.1	0.45 0.22	0.75 -0.35	0.77 -0.39	0.74 -0.34	0.49 0.01	0.68 -0.23	0.61 -0.16	0.53 -0.09	0.59 -0.12	0.94 -0.79
Asian	0.84 -0.62	0.04 0.95	0.58 -0.13	0.72 -0.36	0.78 -0.52	0.41 0.18	0.25 0.4	0.41 0.15	0.68 -0.26	0.48 0.04	0.25 0.41	0.2 0.51	0.47 0.06
Weight	<0.01 1.73	0.02 0.89	0.02 0.91	0.29 0.25	<0.01 1.03	<0.01 1.51	<0.01 1.39	<0.01 1.28	0.01 1.09	<0.01 1.23	<0.01 1.19	0.01 1.07	<0.01 1.2
Skin-Tone	0.1 0.72	0.28 0.32	0.65 -0.21	0.03 0.92	0.56 -0.09	0.14 0.58	0.39 -0.12	0.36 0.2	0.26 0.35	0.46 0.05	0.41 0.12	0.34 0.21	0.48 0.02
Disability	0.63 -0.26	0.5 0.0	0.13 0.89	0.53 -0.04	0.66 -0.19	0.41 0.2	0.03 1.42	0.89 -0.89	0.61 -0.32	0.84 -0.8	0.14 0.84	0.13 0.91	0.63 -0.39
President - Kennedy vs. Trump	0.49 0.02	0.9 -0.76	0.93 -0.86	0.48 0.27	0.69 -0.3	0.86 -0.63	0.99 -1.23	0.67 -0.22	0.38 0.21	0.57 -0.1	0.98 -1.12	0.98 -1.18	0.58 -0.12
President - B. Clinton vs. Trump	0.36 0.28	0.97 -1.04	0.98 -1.14	0.96 -1.04	0.99 -1.18	0.99 -1.34	0.95 -0.92	0.99 -1.12	0.93 -0.86	0.99 -1.35	0.98 -1.15	0.96 -1.03	0.99 -1.25
President - Bush vs. Trump	0.66 -0.21	0.77 -0.44	0.62 -0.2	0.18 0.76	0.48 0.04	0.67 -0.25	0.43 0.11	0.53 -0.08	0.62 -0.2	0.77 -0.45	0.65 -0.23	0.35 0.23	0.78 -0.47
President - Lincoln vs. Trump	0.19 0.6	0.47 0.05	0.5 0.01	0.49 0.3	0.01 1.23	0.11 0.71	0.91 -0.78	0.13 0.64	0.01 1.43	0.04 1.02	0.47 0.04	0.66 -0.25	0.09 0.81
Religion	0.88 -0.67	0.54 -0.05	0.28 0.31	0.65 -0.28	0.34 0.23	0.44 0.09	0.41 0.13	0.49 0.02	0.52 -0.03	0.45 0.06	0.38 0.17	0.35 0.22	0.32 0.25
Sexuality	0.46 0.05	0.46 0.05	0.33 0.21	0.44 0.06	0.47 0.02	0.27 0.31	0.27 0.3	0.38 0.14	0.17 0.46	0.28 0.28	0.57 -0.09	0.42 0.11	0.24 0.35
Race	0.36 0.23	0.12 0.73	0.03 1.1	0.02 1.08	0.33 0.27	0.15 0.64	<0.01 1.41	0.21 0.48	0.13 0.68	0.86 -0.64	0.01 1.25	0.12 0.73	0.88 -0.7
Arab-Muslim	0.05 0.73	0.03 0.84	0.01 0.92	<0.01 1.15	0.03 0.81	0.04 0.78	0.02 0.9	0.02 0.94	0.08 0.64	0.03 0.84	0.08 0.64	0.03 0.84	0.03 0.85
Age	0.5 -0.01	0.94 -0.91	0.91 -0.8	0.12 0.7	0.9 -0.74	0.99 -1.39	0.71 -0.35	0.99 -1.19	0.98 -1.11	0.99 -1.65	0.54 -0.07	0.81 -0.52	0.99 -1.71
Gender-Science	0.71 -0.17	0.39 0.06	0.48 0.01	0.79 -0.19	0.62 -0.07	0.21 0.18	0.13 0.25	0.41 0.05	0.13 0.25	0.73 -0.14	0.46 0.02	0.07 0.32	0.12 0.26
Gender-Career	0.01 0.58	0.74 -0.15	0.6 -0.05	0.9 -0.28	0.86 -0.24	0.56 -0.04	0.88 -0.26	0.95 -0.37	0.99 -0.6	0.98 -0.47	0.28 0.13	0.98 0.13	0.98 -0.43
Intersectional-Gender-Science-MF	0.83 -0.25	0.99 -0.74	0.99 -0.72	0.93 -0.33	0.99 -0.67	0.99 -0.58	0.99 -0.55	0.99 -0.62	0.99 -0.62	0.99 -0.63	0.99 -0.62	0.99 -0.55	0.99 -0.68
Intersectional-Gender-Science-WMBM	0.56 -0.05	0.13 0.36	0.26 0.21	0.01 0.7	0.1 0.42	0.37 0.12	0.14 0.35	0.29 0.19	0.43 0.05	0.46 0.03	0.31 0.17	0.16 0.32	0.26 0.21
Intersectional-Gender-Science-WMBF	0.89 -0.42	0.78 -0.26	0.77 -0.24	0.03 0.61	0.48 0.01	0.68 -0.16	0.29 0.18	0.62 -0.11	0.8 -0.29	0.81 -0.31	0.58 -0.06	0.41 0.08	0.71 -0.19
Intersectional-Gender-Science-WMWF	0.71 -0.22	0.99 -0.82	0.99 -0.97	0.97 -0.57	0.99 -0.9	0.99 -0.74	0.99 -0.85	0.99 -0.8	0.99 -0.7	0.99 -0.75	0.99 -0.88	0.99 -0.75	0.99 -0.82
Intersectional-Gender-Science-BMBF	0.8 -0.28	0.99 -0.7	0.94 -0.49	0.74 -0.21	0.93 -0.48	0.89 -0.39	0.77 -0.23	0.91 -0.42	0.96 -0.56	0.94 -0.48	0.84 -0.32	0.86 -0.34	0.96 -0.55
Intersectional-Gender-Science-BMWF	0.6 -0.09	0.99 -1.2	0.99 -1.19	0.99 -1.21	0.99 -1.41	0.99 -1.23	0.99 -1.32	0.99 -1.3	0.99 -1.15	0.99 -1.03	0.99 -1.2	0.99 -1.24	0.99 -1.19
Intersectional-Gender-Career-MF	0.03 0.41	0.02 0.46	0.01 0.55	0.42 0.05	0.14 0.25	<0.01 0.69	0.15 0.23	0.01 0.49	0.87 -0.25	0.37 0.07	<0.01 0.79	<0.01 0.68	0.13 0.25
Intersectional-Gender-Career-WMBM	0.26 0.21	0.03 0.58	0.03 0.63	0.03 0.62	0.47 0.02	0.23 0.24	0.01 0.73	0.12 0.38	0.8 -0.27	0.35 0.13	0.04 0.54	0.02 0.65	0.15 0.34
Intersectional-Gender-Career-WMBF	0.19 0.29	0.74 -0.21	0.63 -0.11	0.99 -0.69	0.16 0.32	0.1 0.41	0.85 -0.33	0.44 0.05	0.6 -0.08	0.75 -0.22	0.06 0.49	0.34 0.14	0.66 -0.13
Intersectional-Gender-Career-WMWF	0.11 0.39	0.05 0.51	0.05 0.52	0.34 0.14	0.35 0.13	0.01 0.74	0.58 -0.07	0.05 0.52	0.67 -0.15	0.22 0.25	0.03 0.59	0.03 0.59	0.23 0.25
Intersectional-Gender-Career-BMBF	0.08 0.44	0.08 0.46	0.02 0.64	0.52 -0.02	0.13 0.36	0.02 0.66	0.04 0.56	0.07 0.48	0.85 -0.34	0.62 -0.1	<0.01 0.98	0.01 0.8	0.19 0.27
Intersectional-Gender-Career-BMWF	0.05 0.53	<0.01 1.06	<0.01 1.1	<0.01 0.79	0.31 0.16	<0.01 0.98	<0.01 0.84	<0.01 1.0	0.93 -0.47	0.09 0.43	<0.01 1.06	<0.01 1.15	0.01 0.77
Intersectional-Valence-BW	0.79 -0.18	0.02 0.45	<0.01 0.67	<0.01 1.04	<0.01 0.74	<0.01 0.57	<0.01 0.73	<0.01 0.68	0.01 0.48	0.03 0.43	0.08 0.32	<0.01 0.67	0.02 0.45
Intersectional-Valence-WMBM	0.73 -0.21	0.21 0.26	0.06 0.5	0.01 0.81	0.07 0.46	0.21 0.27	0.13 0.35	0.1 0.41	0.4 0.09	0.4 0.09	0.35 0.13	0.12 0.39	0.32 0.15
Intersectional-Valence-WMBF	0.75 -0.23	0.68 -0.16	0.32 0.15	0.01 0.8	0.26 0.2	0.51 -0.01	0.15 0.33	0.43 0.06	0.44 0.05	0.63 -0.11	0.51 -0.02	0.23 0.23	0.48 0.02
Intersectional-Valence-WMWF	0.35 0.15	0.01 0.72	0.02 0.67	0.03 0.58	<0.01 0.84	<0.01 0.81	0.01 0.78	<0.01 0.83	<0.01 0.79	<0.01 0.78	0.07 0.47	0.01 0.71	0.02 0.65
Intersectional-Valence-WFBM	0.64 -0.12	<0.01 1.1	<0.01 1.14	<0.01 1.26	<0.01 1.26	<0.01 1.23	<0.01 1.1	<0.01 1.31	<0.01 1.01	<0.01 1.15	0.02 0.64	<0.01 1.08	<0.01 0.94
Intersectional-Valence-BFBM	0.45 0.04	0.05 0.51	0.09 0.43	0.37 0.1	0.19 0.28	0.13 0.36	0.41 0.07	0.08 0.45	0.44 0.04	0.2 0.28	0.3 0.17	0.27 0.2	0.3 0.17
Intersectional-Valence-WFBF	0.68 -0.15	0.01 0.77	<0.01 0.92	<0.01 1.34	<0.01 1.06	<0.01 1.0	<0.01 1.15	<0.01 1.15	<0.01 0.96	<0.01 1.01	0.04 0.54	<0.01 1.01	<0.01 0.85
Intersectional-Valence-FM	0.35 0.09	<0.01 0.61	0.01 0.53	0.13 0.25	0.01 0.54	<0.01 0.59	0.03 0.43	<0.01 0.62	0.03 0.43	0.01 0.56	0.07 0.33	0.02 0.44	0.03 0.43

Table 6. Results for the detection of biases based on the 5th block embeddings of 13 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.99 -0.6	0.99 -0.56	<0.01 0.76	0.99 -0.58	<0.01 1.0	<0.01 1.42	<0.01 1.57	<0.01 1.18	<0.01 1.41	0.37 0.08	<0.01 1.73	<0.01 1.78	<0.01 1.77
Weapon	0.01 1.41	0.55 -0.08	0.96 -1.04	0.14 0.66	0.97 -1.11	0.94 -0.92	0.95 -0.91	0.97 -1.08	0.69 -0.24	0.59 -0.14	0.14 0.63	0.15 0.65	0.95 -0.95
Weapon (Modern)	0.07 0.86	0.77 -0.44	0.5 -0.0	0.11 0.71	0.88 -0.68	0.83 -0.6	0.93 -0.86	0.99 -1.36	0.85 -0.59	0.99 -1.35	0.05 0.95	0.15 0.62	0.99 -1.27
Native	0.89 -0.62	0.41 0.12	0.03 0.9	0.62 -0.15	0.97 -0.94	0.84 -0.53	0.93 -0.73	0.34 0.2	0.98 -1.03	0.55 -0.08	0.66 -0.21	0.44 0.08	0.99 -1.29
Asian	0.87 -0.67	0.02 1.2	0.51 -0.01	0.55 -0.09	0.61 -0.19	0.44 0.11	0.53 -0.05	0.74 -0.37	0.86 -0.67	0.58 -0.11	0.96 -1.0	0.9 -0.75	0.38 0.21
Weight	<0.01 1.66	0.02 0.9	0.88 -0.54	0.29 0.28	0.99 -1.53	0.99 -0.98	0.07 0.65	0.91 -0.6	0.95 -0.71	0.28 0.29	0.94 -0.71	0.01 0.97	0.41 0.11
Skin-Tone	0.21 0.52	0.76 -0.4	0.18 0.51	0.05 0.86	0.46 0.05	0.35 0.22	0.84 -0.54	0.14 0.6	0.72 -0.36	0.59 -0.22	0.41 0.13	0.92 -0.75	0.96 -0.92
Disability	0.7 -0.35	0.16 0.81	0.07 1.11	0.29 0.49	0.87 -0.86	0.77 -0.53	0.01 1.45	0.87 -0.84	0.2 0.66	0.9 -1.07	0.67 -0.3	0.59 -0.15	0.47 0.08
President - Kennedy vs. Trump	0.47 0.04	0.8 -0.5	0.87 -0.66	0.41 0.15	0.98 -1.12	0.93 -0.87	0.95 -0.92	0.99 -1.56	0.99 -1.47	0.67 -0.23	0.99 -1.4	0.99 -1.44	0.93 -0.88
President - B. Clinton vs. Trump	0.37 0.28	0.39 0.16	0.98 -1.11	0.77 -0.46	0.99 -1.25	0.99 -1.28	0.99 -1.18	0.99 -1.49	0.99 -1.24	0.99 -1.0	0.99 -1.52	0.99 -1.24	0.99 -1.43
President - Bush vs. Trump	0.67 -0.23	0.71 -0.34	0.83 -0.54	0.2 0.51	0.72 -0.36	0.91 -0.78	0.59 -0.15	0.9 -0.77	0.95 -0.94	0.6 -0.14	0.89 -0.75	0.78 -0.52	0.94 -0.93
President - Lincoln vs. Trump	0.21 0.57	0.02 1.2	0.22 0.46	0.1 0.78	0.98 -1.21	0.25 0.4	0.94 -0.92	0.81 -0.53	0.71 -0.34	0.2 0.51	0.89 -0.73	0.85 -0.65	0.51 -0.01
Religion	0.88 -0.67	0.49 0.01	0.38 0.17	0.64 -0.2	0.43 0.31	0.62 0.06	0.23 0.09	0.42 -0.17	0.42 0.1	0.42 0.14	0.4 0.14	0.37 0.19	0.29 0.32
Sexuality	0.43 0.08	0.32 0.23	0.08 0.67	0.55 -0.06	0.4 0.12	0.26 0.32	0.29 0.29	0.15 0.49	0.15 0.53	0.45 0.06	0.48 0.03	0.45 0.07	0.61 -0.15
Race	0.38 0.19	0.02 1.18	0.28 0.34	0.01 1.24	0.03 1.08	0.19 0.52	0.89 -0.74	0.04 0.98	0.18 0.58	0.97 -1.17	0.77 -0.45	0.1 0.79	0.94 -0.92
Arab-Muslim	0.07 0.69	0.09 0.59	0.17 0.43	0.01 1.06	0.01 1.05	0.01 0.97	0.09 0.6	0.33 0.19	0.35 0.18	0.01 1.12	0.05 0.72	0.1 0.59	0.1 0.58
Age	0.52 -0.04	0.99 -1.55	0.99 -1.43	0.95 -0.96	0.11 0.74	0.99 -1.39	0.98 -1.07	0.54 -0.06	0.98 -1.17	0.99 -1.69	0.81 -0.52	0.52 -0.02	0.99 -1.75
Gender-Science	0.78 -0.3	0.67 -0.09	0.01 0.54	0.7 -0.12	0.05 0.37	0.79 -0.18	0.04 0.39	0.79 -0.18	0.61 -0.07	0.74 -0.14	0.43 0.04	0.01 0.51	<0.01 0.82
Gender-Career	<0.01 0.61	0.2 0.19	0.99 -0.64	0.95 -0.36	0.65 -0.09	0.25 0.15	0.54 -0.02	0.01 0.5	0.02 0.46	0.96 -0.38	0.34 0.09	0.91 -0.3	0.09 0.3
Intersectional-Gender-Science-MF	0.91 -0.36	0.99 -0.65	0.02 0.44	0.99 -0.56	0.86 -0.24	0.98 -0.54	0.99 -0.54	0.71 -0.12	0.77 -0.18	0.99 -0.66	0.86 -0.24	0.36 0.08	<0.01 0.61
Intersectional-Gender-Science-WMBM	0.45 0.04	0.28 0.19	0.11 0.39	0.01 0.71	0.36 0.1	0.23 0.24	0.1 0.4	<0.01 0.9	<0.01 0.87	0.08 0.44	<0.01 0.8	0.03 0.6	0.18 0.29
Intersectional-Gender-Science-WMBF	0.86 -0.44	0.81 -0.28	0.07 0.48	0.06 0.5	0.3 0.17	0.47 0.02	0.45 0.04	<0.01 0.97	<0.01 0.85	0.42 0.07	0.05 0.51	0.04 0.56	<0.01 1.0
Intersectional-Gender-Science-WMWF	0.83 -0.36	0.99 -0.9	<0.01 0.79	0.99 -0.83	0.94 -0.49	0.99 -0.68	0.98 -0.63	0.85 -0.34	0.9 -0.39	0.99 -0.85	0.75 -0.21	0.34 0.14	0.07 0.47
Intersectional-Gender-Science-BMBF	0.87 -0.38	0.92 -0.44	0.38 0.1	0.91 -0.43	0.37 0.1	0.8 -0.27	0.94 -0.5	0.32 0.15	0.48 0.02	0.94 -0.52	0.83 -0.3	0.47 0.01	<0.01 0.85
Intersectional-Gender-Science-BMWF	0.81 -0.29	0.99 -1.03	0.1 0.42	0.99 -1.37	0.98 -0.66	0.99 -0.97	0.99 -1.17	0.99 -1.04	0.99 -1.24	0.99 -1.42	0.99 -1.03	0.93 -0.47	0.23 0.24
Intersectional-Gender-Career-MF	0.03 0.4	0.11 0.28	0.97 -0.43	0.15 0.23	<0.01 0.89	<0.01 0.73	<0.01 0.74	<0.01 0.98	<0.01 0.91	<0.01 0.77	<0.01 0.79	<0.01 0.62	<0.01 1.05
Intersectional-Gender-Career-WMBM	0.28 0.19	0.1 0.4	0.04 0.55	<0.01 0.89	0.09 0.42	0.09 0.43	0.01 0.72	0.09 0.42	0.11 0.4	0.04 0.55	0.01 0.69	0.01 0.78	0.17 0.3
Intersectional-Gender-Career-WMBF	0.17 0.3	0.49 0.0	0.98 -0.65	0.99 -0.86	0.08 0.45	0.15 0.34	0.53 -0.03	0.03 0.62	0.02 0.61	0.23 0.24	0.38 0.11	0.58 -0.06	0.02 0.66
Intersectional-Gender-Career-WMWF	0.13 0.37	0.32 0.15	0.99 -0.75	0.14 0.36	<0.01 0.97	0.01 0.71	<0.01 0.79	<0.01 0.94	<0.01 0.81	0.02 0.66	0.01 0.71	0.12 0.38	<0.01 1.08
Intersectional-Gender-Career-BMBF	0.08 0.43	0.12 0.38	0.59 -0.07	0.25 0.21	<0.01 0.85	0.01 0.8	0.01 0.81	<0.01 1.03	<0.01 1.02	<0.01 0.93	<0.01 0.99	<0.01 1.01	<0.01 1.07
Intersectional-Gender-Career-BMWF	0.06 0.5	0.05 0.53	0.79 -0.25	<0.01 1.19	<0.01 1.34	<0.01 1.09	<0.01 1.47	<0.01 1.35	<0.01 1.18	<0.01 1.35	<0.01 1.42	<0.01 1.44	<0.01 1.4
Intersectional-Valence-BW	0.82 -0.21	0.03 0.42	0.75 -0.15	<0.01 1.15	0.12 0.27	0.49 0.01	0.38 0.08	<0.01 0.59	<0.01 0.64	0.01 0.55	0.14 0.24	0.02 0.47	0.99 -0.52
Intersectional-Valence-WMBM	0.76 -0.24	0.13 0.36	0.83 -0.3	<0.01 0.82	0.1 0.41	0.63 -0.1	0.59 -0.07	0.16 0.32	0.04 0.55	0.19 0.28	0.36 0.12	0.1 0.4	0.52 -0.03
Intersectional-Valence-WMBF	0.81 -0.28	0.5 -0.0	0.24 0.24	<0.01 0.88	0.43 0.06	0.87 -0.36	0.79 -0.25	0.02 0.63	0.01 0.71	0.67 -0.15	0.4 0.09	0.26 0.21	0.79 -0.26
Intersectional-Valence-WMWF	0.32 0.18	0.08 0.45	0.71 -0.17	0.01 0.71	0.47 0.03	0.07 0.48	0.08 0.46	0.21 0.26	0.51 -0.01	<0.01 0.91	0.21 0.27	0.11 0.4	0.99 -0.69
Intersectional-Valence-WFBM	0.64 -0.12	0.01 0.81	0.94 -0.49	<0.01 1.42	0.06 0.49	0.12 0.39	0.14 0.35	0.03 0.57	0.04 0.56	<0.01 1.32	0.11 0.38	0.02 0.66	0.99 -0.81
Intersectional-Valence-BFBM	0.41 0.07	0.11 0.4	0.96 -0.56	0.49 0.0	0.12 0.38	0.21 0.26	0.32 0.15	0.83 -0.3	0.65 -0.13	0.05 0.52	0.44 0.05	0.26 0.21	0.18 0.29
Intersectional-Valence-WFBF	0.72 -0.19	0.06 0.5	0.41 0.07	<0.01 1.53	0.38 0.1	0.32 0.16	0.24 0.24	<0.01 0.86	0.01 0.73	<0.01 1.09	0.12 0.39	0.04 0.54	0.99 -0.98
Intersectional-Valence-FM	0.29 0.12	0.03 0.42	0.95 -0.37	0.1 0.29	0.18 0.21	0.05 0.37	0.09 0.3	0.49 0.0	0.59 -0.06	<0.01 0.7	0.23 0.16	0.09 0.3	0.89 -0.27

Table 7. Results for the detection of biases based on the GAP-layer embeddings of 13 models. Each cell contains two numbers corresponding to the test results: the p -value (at the top) and the d -value (on the bottom). Cells are colored according to the p -values they contain.

	Random	Jigsaw	Rotation	Relative location	ClusterFit	ODC	SwAV	NPID	MoCo_v1	MoCo_v2	SimCLR	BYOL	Supervised
Insect-Flower	0.97 -0.54	0.99 -0.81	0.03 0.45	0.99 -0.57	<0.01 0.81	<0.01 1.55	<0.01 1.5	<0.01 0.79	<0.01 1.22	<0.01 1.49	<0.01 1.69	<0.01 1.81	<0.01 1.73
Weapon	0.35 0.23	0.81 -0.51	0.13 0.7	0.08 0.85	0.93 -0.85	0.23 0.45	0.85 -0.61	0.71 -0.33	0.54 -0.06	0.9 -0.78	0.97 -1.05	0.17 0.56	0.99 -1.28
Weapon (Modern)	0.14 0.68	0.67 -0.27	0.47 0.05	0.05 0.93	0.66 -0.26	0.55 -0.08	0.92 -0.86	0.99 -1.44	0.83 -0.57	0.99 -1.24	0.37 0.21	0.68 -0.28	0.99 -1.36
Native	0.76 -0.37	0.74 -0.35	0.03 0.92	0.85 -0.53	0.99 -1.37	0.98 -1.01	0.91 -0.69	0.92 -0.7	0.99 -1.44	0.99 -1.03	0.99 -1.34	0.63 -0.17	0.99 -1.51
Asian	0.72 -0.35	0.48 0.02	0.55 -0.08	0.52 -0.04	0.91 -0.81	0.99 -1.47	0.89 -0.72	0.99 -1.34	0.98 -1.18	0.71 -0.35	0.94 -0.91	0.99 -1.19	0.13 0.7
Weight	<0.01 1.88	<0.01 1.45	0.82 -0.43	0.99 -1.45	0.99 -1.28	0.83 -0.44	0.03 0.86	0.66 -0.2	0.36 0.17	0.12 0.54	0.75 -0.31	<0.01 1.21	0.31 0.22
Skin-Tone	<0.01 1.4	0.88 -0.65	0.23 0.4	0.11 0.69	0.68 -0.27	0.38 0.16	0.84 -0.56	0.2 0.47	0.69 -0.28	0.79 -0.45	0.31 0.22	0.9 -0.7	0.91 -0.73
Disability	0.44 0.23	0.01 1.54	0.01 1.54	0.03 1.21	0.63 -0.33	0.43 0.12	0.14 0.8	0.61 -0.17	0.13 0.87	0.56 -0.14	0.44 0.39	0.8 -0.61	0.41 0.16
President - Kennedy vs. Trump	0.5 0.02	0.43 0.12	0.63 -0.23	0.43 0.12	0.98 -1.08	0.94 -0.92	0.89 -0.74	0.99 -1.33	0.99 -1.26	0.92 -0.83	0.88 -0.68	0.92 -0.85	0.9 -0.76
President - B. Clinton vs. Trump	0.44 0.08	0.04 1.01	0.22 0.46	0.17 0.62	0.99 -1.24	0.85 -0.63	0.97 -1.06	0.99 -1.43	0.99 -1.4	0.84 -0.6	0.99 -1.29	0.9 -0.79	0.96 -1.03
President - Bush vs. Trump	0.44 0.08	0.23 0.52	0.21 0.47	0.02 0.94	0.79 -0.49	0.35 0.25	0.47 0.06	0.73 -0.36	0.81 -0.53	0.42 0.12	0.43 0.12	0.56 -0.11	0.94 -0.9
President - Lincoln vs. Trump	0.05 0.98	0.01 1.42	0.1 0.74	<0.01 1.27	0.98 -1.23	0.38 0.18	0.92 -0.86	0.81 -0.54	0.73 -0.38	0.42 0.11	0.46 0.07	0.51 -0.01	0.2 0.49
Religion	0.73 -0.34	0.31 0.28	0.17 0.54	0.54 -0.07	0.32 0.27	0.42 0.11	0.35 0.22	0.34 0.24	0.13 0.61	0.2 0.46	0.21 0.44	0.25 0.38	0.11 0.68
Sexuality	0.24 0.34	0.42 0.1	0.14 0.52	0.64 -0.18	0.28 0.3	0.14 0.55	0.35 0.2	0.17 0.48	0.16 0.48	0.61 -0.13	0.28 0.3	0.33 0.23	0.76 -0.35
Race	0.24 0.44	0.02 1.12	0.33 0.26	0.03 1.06	0.13 0.66	0.14 0.64	0.96 -0.99	0.37 0.23	0.39 0.17	0.92 -0.83	0.91 -0.8	0.87 -0.71	0.82 -0.54
Arab-Muslim	0.2 0.4	0.15 0.46	0.17 0.44	0.01 1.08	<0.01 1.1	0.04 0.78	0.1 0.58	0.47 0.04	0.46 0.04	0.03 0.81	0.1 0.6	0.14 0.48	0.12 0.53
Age	0.66 -0.29	0.99 -1.5	0.99 -1.45	0.95 -0.94	0.11 0.74	0.99 -1.37	0.99 -1.2	0.52 -0.02	0.86 -0.66	0.87 -0.66	0.9 -0.77	0.32 0.29	0.99 -1.35
Gender-Science	0.81 -0.25	0.6 -0.06	0.26 0.15	0.14 0.25	0.14 0.25	0.35 0.1	0.23 0.16	0.78 -0.18	0.72 -0.13	0.9 -0.3	0.48 0.01	0.09 0.3	0.03 0.43
Gender-Career	<0.01 0.62	0.72 -0.13	0.84 -0.23	0.54 -0.03	0.12 0.26	0.05 0.37	0.11 0.28	<0.01 0.83	<0.01 0.79	0.36 0.08	<0.01 0.71	0.05 0.37	<0.01 0.69
Intersectional-Gender-Science-MF	0.75 -0.17	0.87 -0.25	0.17 0.22	0.91 -0.3	0.78 -0.18	0.98 -0.44	0.97 -0.43	0.7 -0.12	0.92 -0.32	0.98 -0.45	0.9 -0.28	0.37 0.07	0.01 0.51
Intersectional-Gender-Science-WMBM	<0.01 0.84	0.6 -0.08	0.1 0.41	<0.01 0.86	0.44 0.06	0.05 0.53	0.01 0.7	<0.01 0.86	<0.01 0.96	<0.01 0.88	<0.01 0.9	<0.01 0.81	0.12 0.37
Intersectional-Gender-Science-WMBF	<0.01 1.02	0.37 0.1	0.1 0.41	<0.01 0.89	0.27 0.2	0.09 0.4	0.11 0.4	<0.01 0.96	<0.01 0.88	0.02 0.65	0.02 0.67	<0.01 0.9	<0.01 0.95
Intersectional-Gender-Science-WMWF	0.99 -0.69	0.99 -0.75	0.05 0.52	0.99 -0.76	0.92 -0.44	0.99 -0.69	0.96 -0.53	0.88 -0.37	0.97 -0.6	0.99 -0.72	0.9 -0.4	0.54 -0.03	0.1 0.42
Intersectional-Gender-Science-BMBF	0.19 0.32	0.3 0.17	0.58 -0.06	0.51 -0.01	0.28 0.18	0.75 -0.22	0.9 -0.41	0.33 0.15	0.64 -0.12	0.82 -0.29	0.77 -0.23	0.25 0.22	0.02 0.65
Intersectional-Gender-Science-BMWF	0.99 -1.39	0.97 -0.61	0.45 0.04	0.99 -1.44	0.97 -0.6	0.99 -1.3	0.99 -1.3	0.99 -1.16	0.99 -1.43	0.99 -1.51	0.99 -1.27	0.99 -0.86	0.45 0.04
Intersectional-Gender-Career-MF	0.08 0.31	0.31 0.11	0.47 0.01	0.09 0.3	<0.01 0.93	<0.01 0.6	<0.01 0.65	<0.01 0.93	<0.01 0.95	<0.01 0.84	<0.01 1.1	<0.01 0.69	<0.01 1.03
Intersectional-Gender-Career-WMBM	0.07 0.48	0.04 0.55	0.05 0.52	<0.01 1.16	0.07 0.46	0.42 0.06	0.03 0.6	0.63 -0.11	0.46 0.03	0.03 0.58	0.04 0.55	0.04 0.56	0.39 0.09
Intersectional-Gender-Career-WMBF	0.84 -0.33	0.67 -0.15	0.81 -0.28	0.99 -1.04	0.06 0.5	0.04 0.56	0.38 0.09	<0.01 0.99	<0.01 0.96	0.02 0.63	0.01 0.79	0.29 0.18	<0.01 0.88
Intersectional-Gender-Career-WMWF	0.06 0.5	0.83 -0.31	0.85 -0.32	0.05 0.51	<0.01 1.15	0.03 0.57	0.02 0.67	<0.01 0.97	<0.01 0.98	0.04 0.57	<0.01 1.06	0.01 0.75	<0.01 1.15
Intersectional-Gender-Career-BMBF	0.28 0.18	0.09 0.43	0.16 0.33	0.15 0.33	<0.01 0.83	0.03 0.62	0.01 0.68	<0.01 0.9	<0.01 0.91	<0.01 1.06	<0.01 1.18	0.01 0.69	<0.01 0.91
Intersectional-Gender-Career-BMWF	<0.01 0.89	0.13 0.37	0.26 0.21	<0.01 1.43	<0.01 1.36	0.02 0.64	<0.01 1.12	<0.01 0.86	<0.01 0.93	<0.01 1.04	<0.01 1.35	<0.01 1.15	<0.01 1.16
Intersectional-Valence-BW	<0.01 0.95	0.01 0.54	0.71 -0.12	<0.01 1.2	0.17 0.22	0.27 0.14	0.17 0.21	<0.01 0.62	<0.01 0.74	<0.01 0.63	0.02 0.45	<0.01 0.83	0.99 -0.49
Intersectional-Valence-WMBM	0.04 0.54	0.26 0.21	0.94 -0.51	<0.01 0.95	0.17 0.3	0.53 -0.02	0.46 0.03	0.15 0.33	0.05 0.53	0.16 0.32	0.31 0.16	0.02 0.68	0.54 -0.03
Intersectional-Valence-WMBF	0.01 0.75	0.22 0.25	0.34 0.14	<0.01 0.96	0.59 -0.07	0.54 -0.03	0.6 -0.08	0.02 0.62	<0.01 0.88	0.43 0.06	0.17 0.31	0.02 0.63	0.94 -0.51
Intersectional-Valence-WMWF	0.01 0.63	0.02 0.64	0.27 0.19	<0.01 0.79	0.3 0.17	0.14 0.36	0.06 0.49	0.15 0.33	0.37 0.1	<0.01 0.81	0.09 0.44	0.05 0.51	0.94 -0.5
Intersectional-Valence-WFBM	<0.01 1.21	<0.01 0.83	0.84 -0.32	<0.01 1.43	0.03 0.57	0.15 0.34	0.07 0.47	0.02 0.62	0.03 0.6	<0.01 1.11	0.03 0.58	<0.01 0.98	0.94 -0.51
Intersectional-Valence-BFBM	0.8 -0.27	0.53 -0.02	0.97 -0.61	0.45 0.04	0.1 0.41	0.48 0.01	0.37 0.11	0.79 -0.27	0.9 -0.42	0.13 0.35	0.65 -0.13	0.27 0.2	0.07 0.46
Intersectional-Valence-WFBF	<0.01 1.36	<0.01 0.95	0.17 0.31	<0.01 1.47	0.36 0.11	0.16 0.33	0.07 0.47	<0.01 0.88	<0.01 0.93	<0.01 1.04	0.01 0.78	<0.01 1.01	0.99 -0.94
Intersectional-Valence-FM	0.23 0.17	0.1 0.29	0.85 -0.24	0.11 0.28	0.09 0.29	0.22 0.17	0.1 0.29	0.4 0.06	0.7 -0.12	<0.01 0.58	0.25 0.15	0.07 0.32	0.56 -0.04

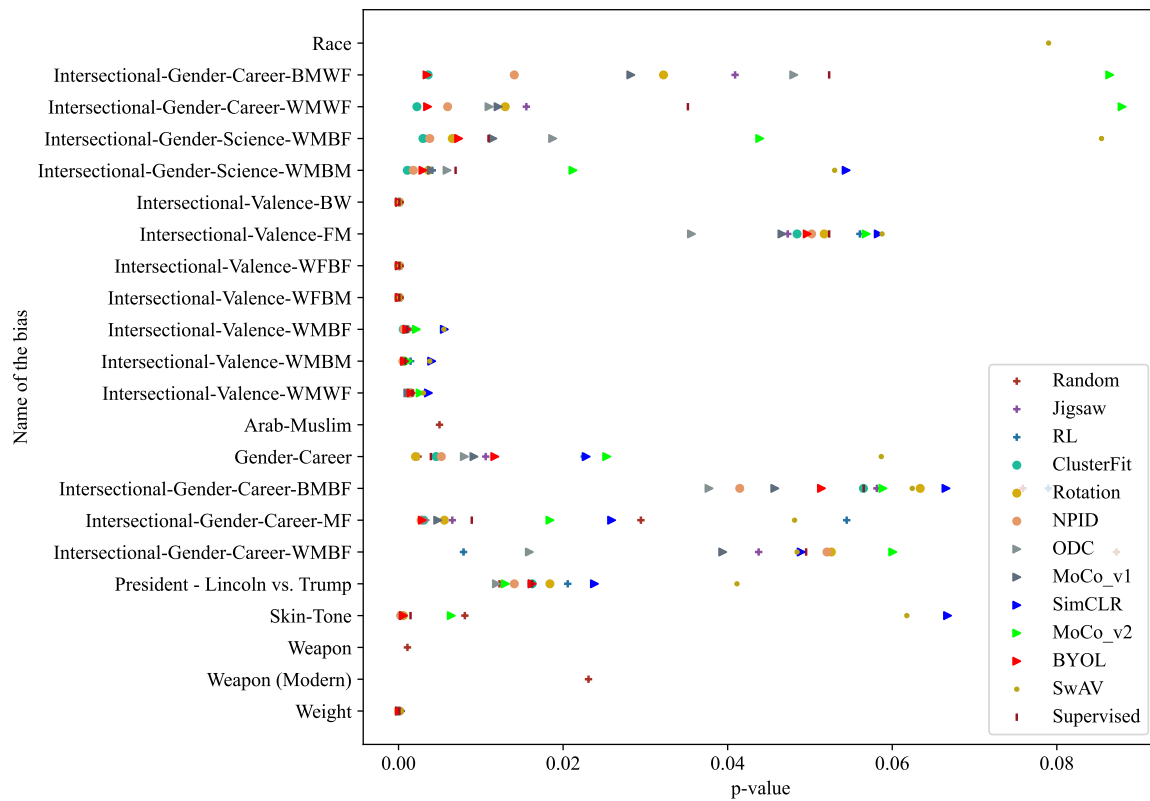


Figure 3. Biases detected in the first ResNet block.

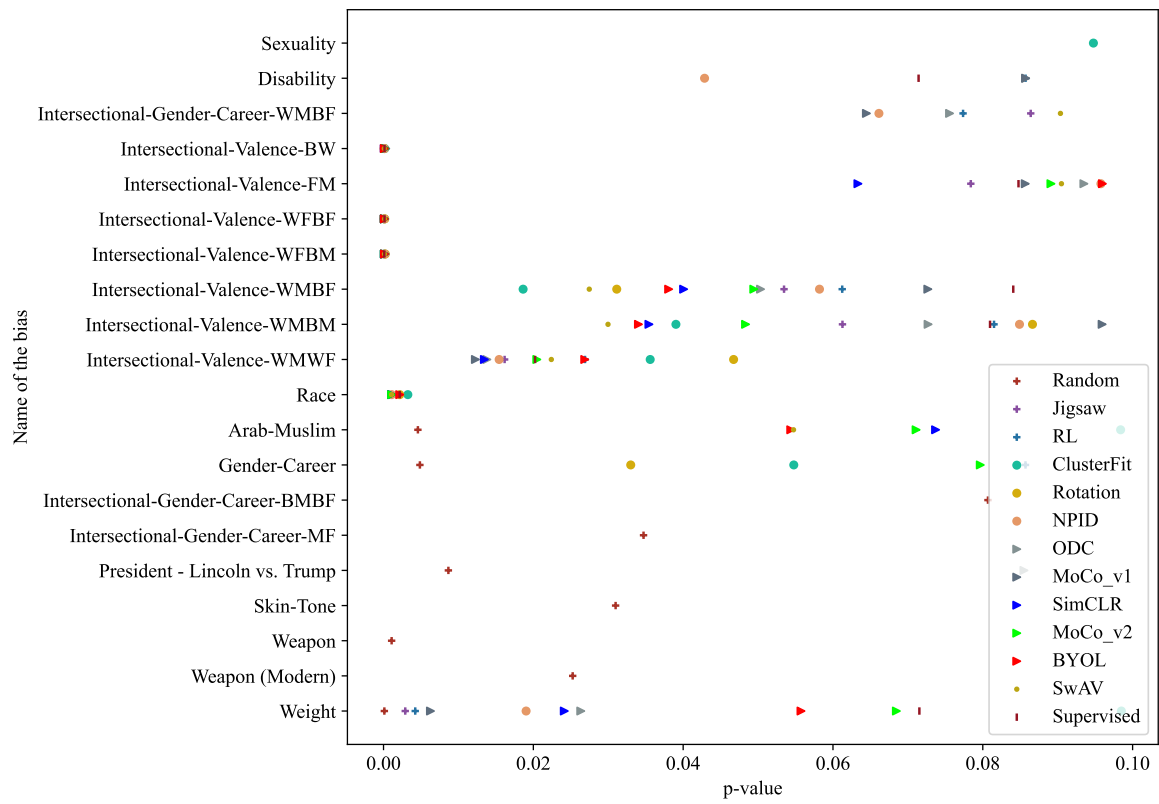


Figure 4. Biases detected in the second ResNet block.

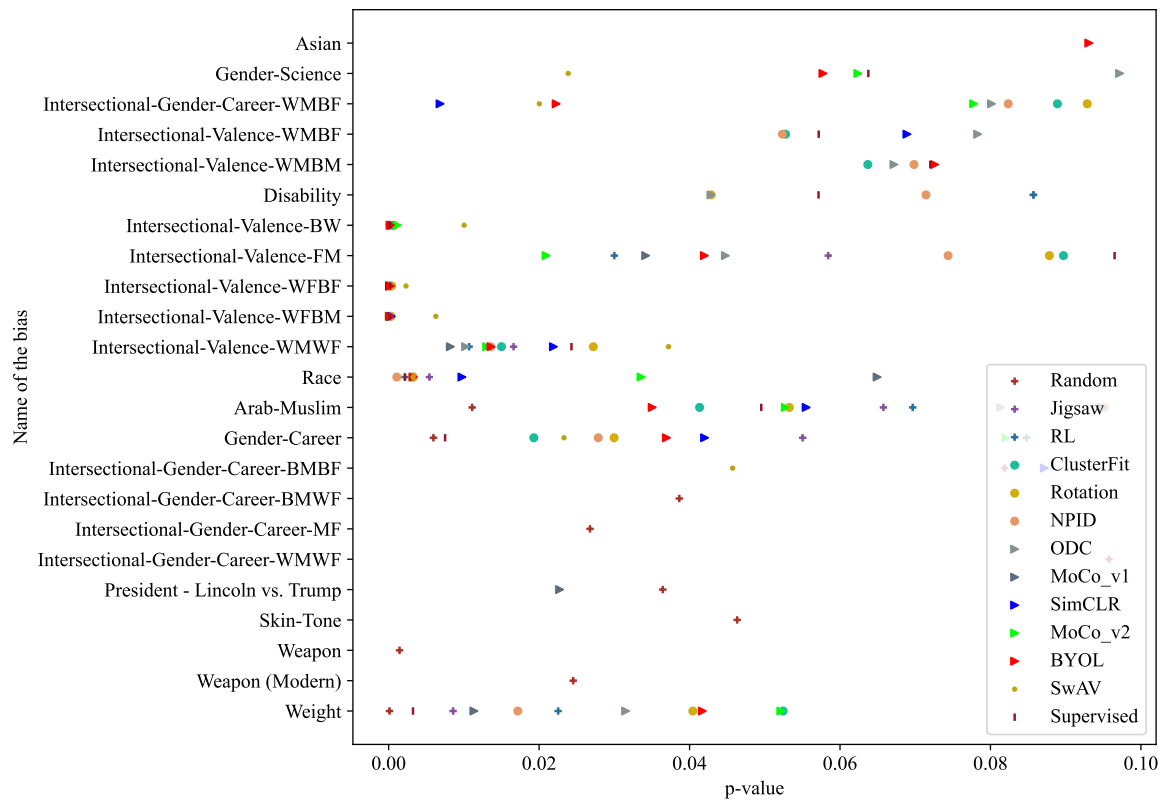


Figure 5. Biases detected in the third ResNet block.

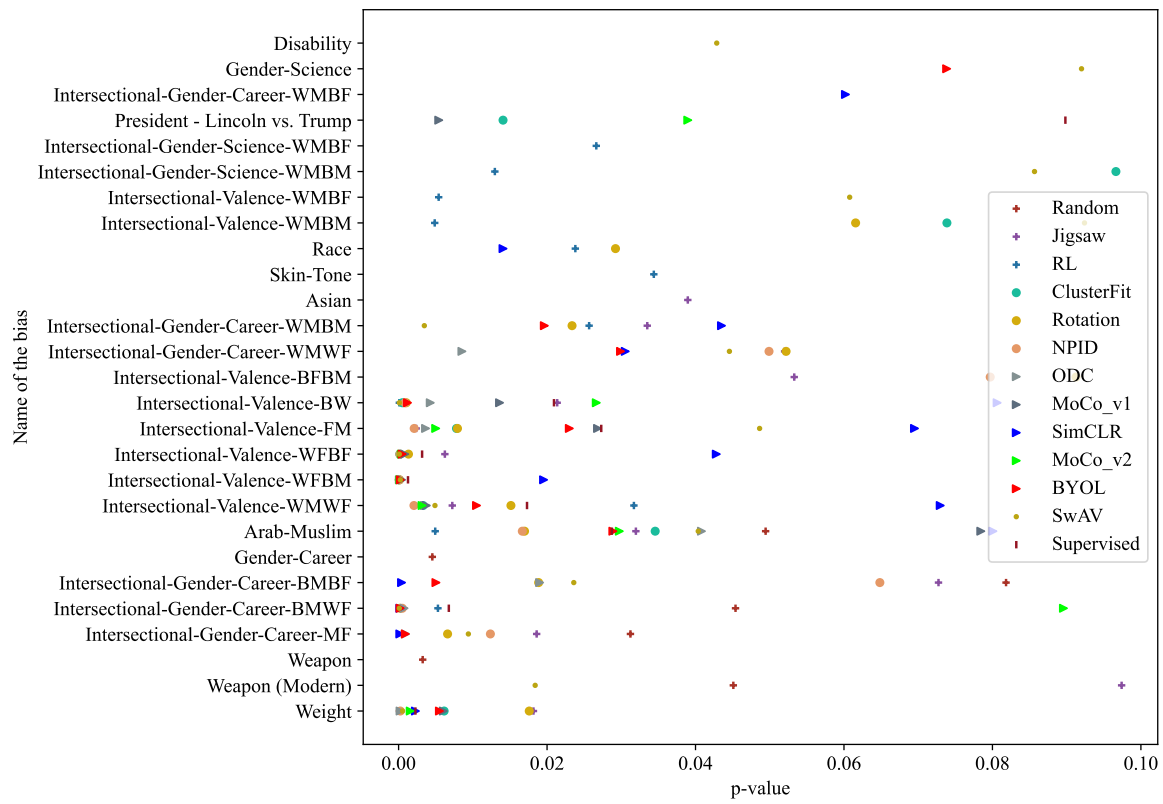


Figure 6. Biases detected in the fourth ResNet block.

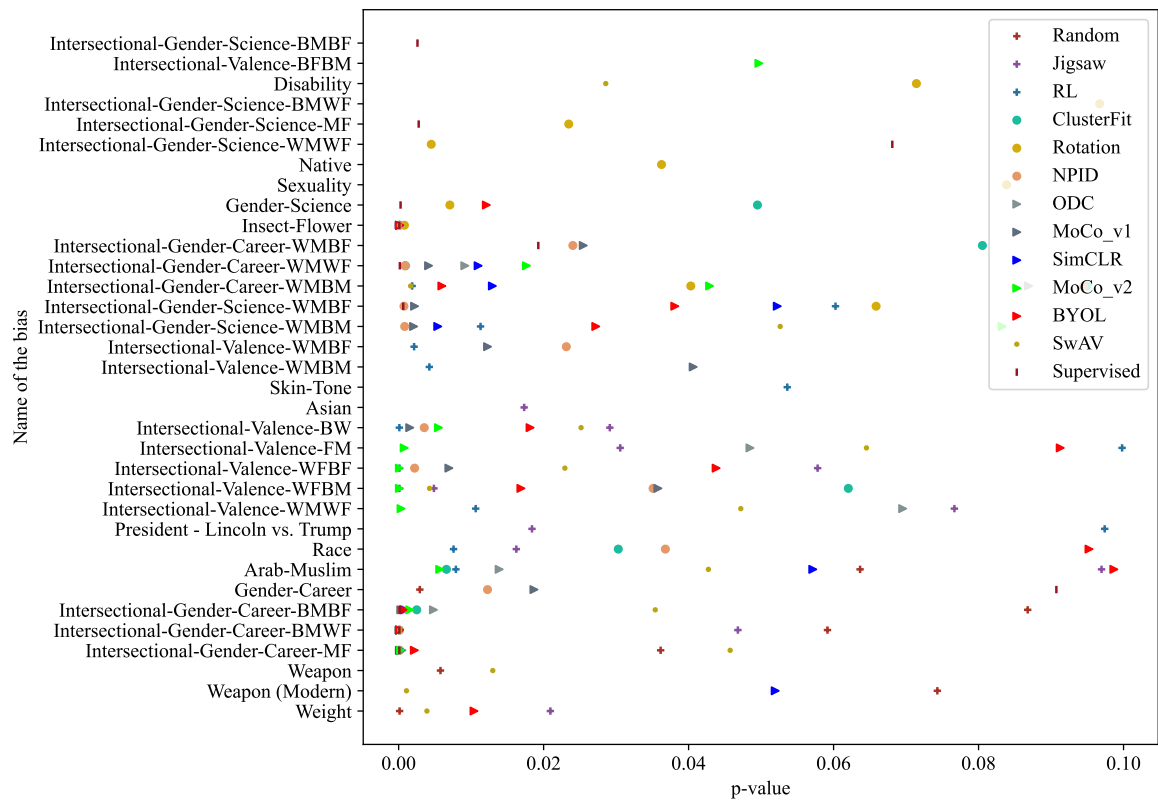


Figure 7. Biases detected in the fifth ResNet block.

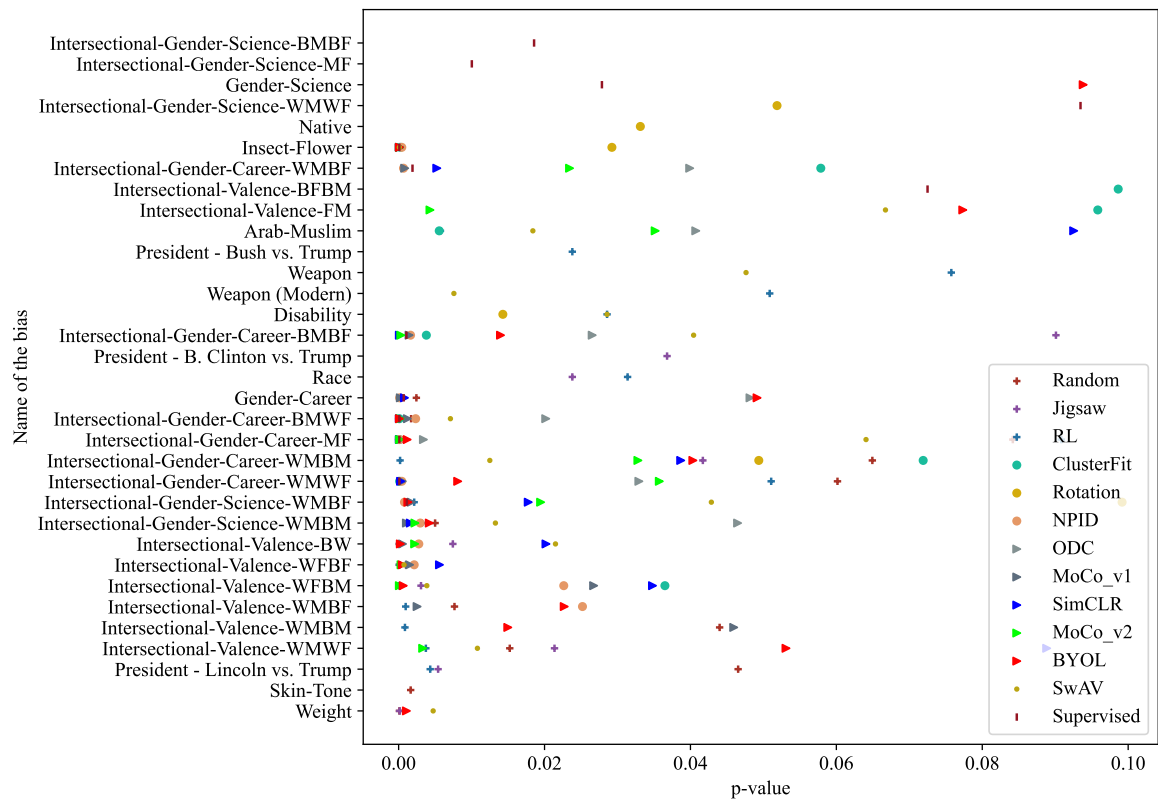


Figure 8. Biases detected in the GAP-layer.

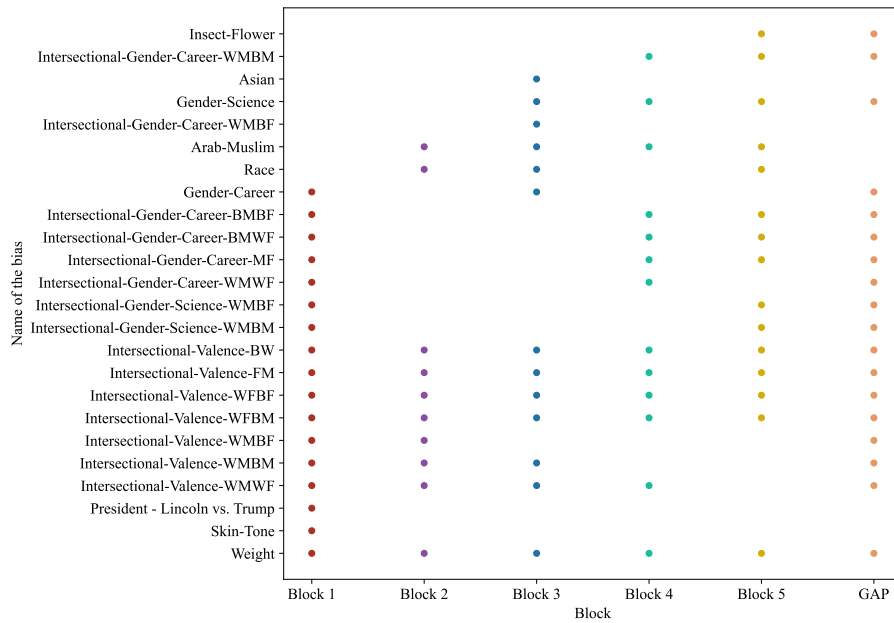


Figure 9. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of BYOL model.

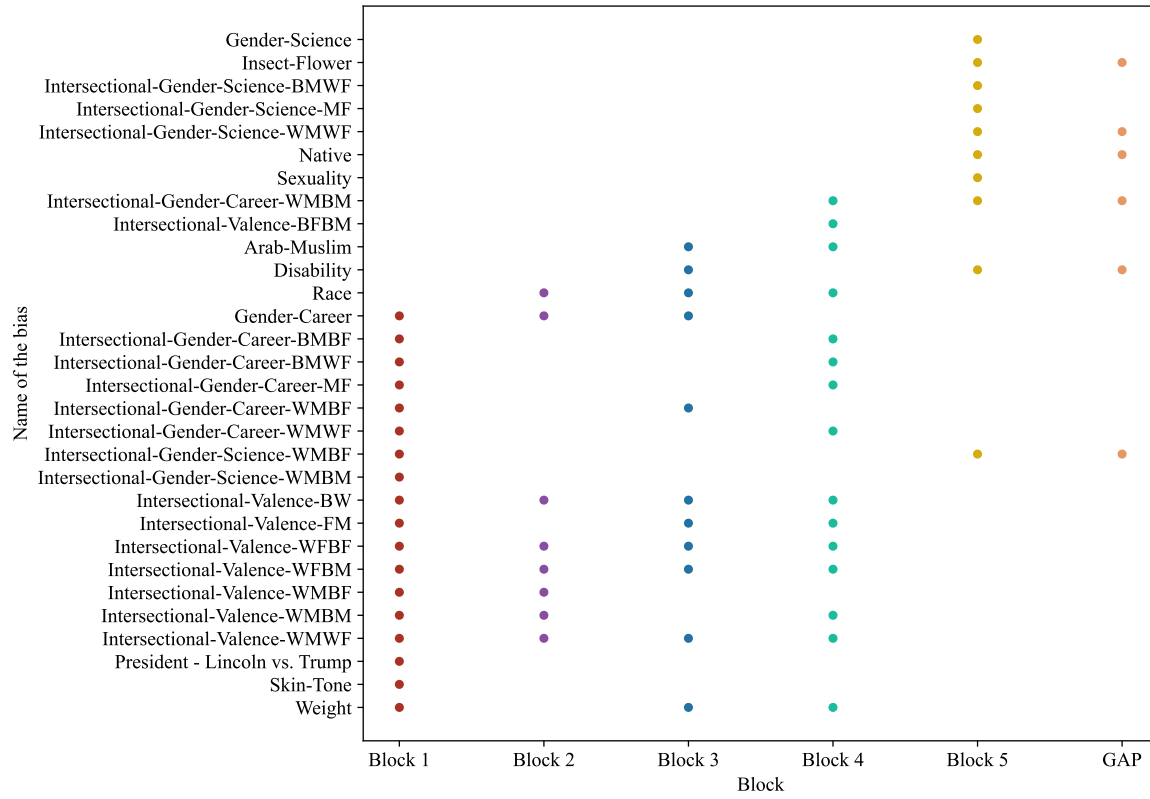


Figure 10. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the rotation prediction model.

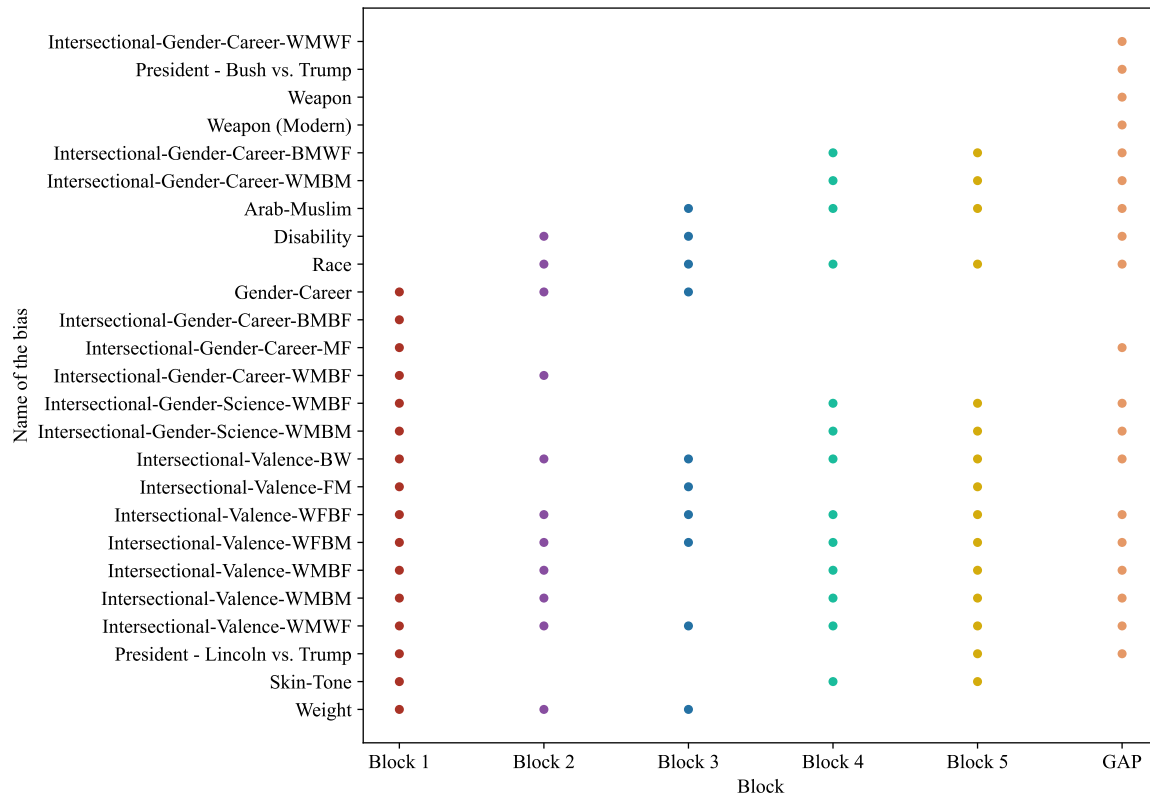


Figure 11. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the relative location model.

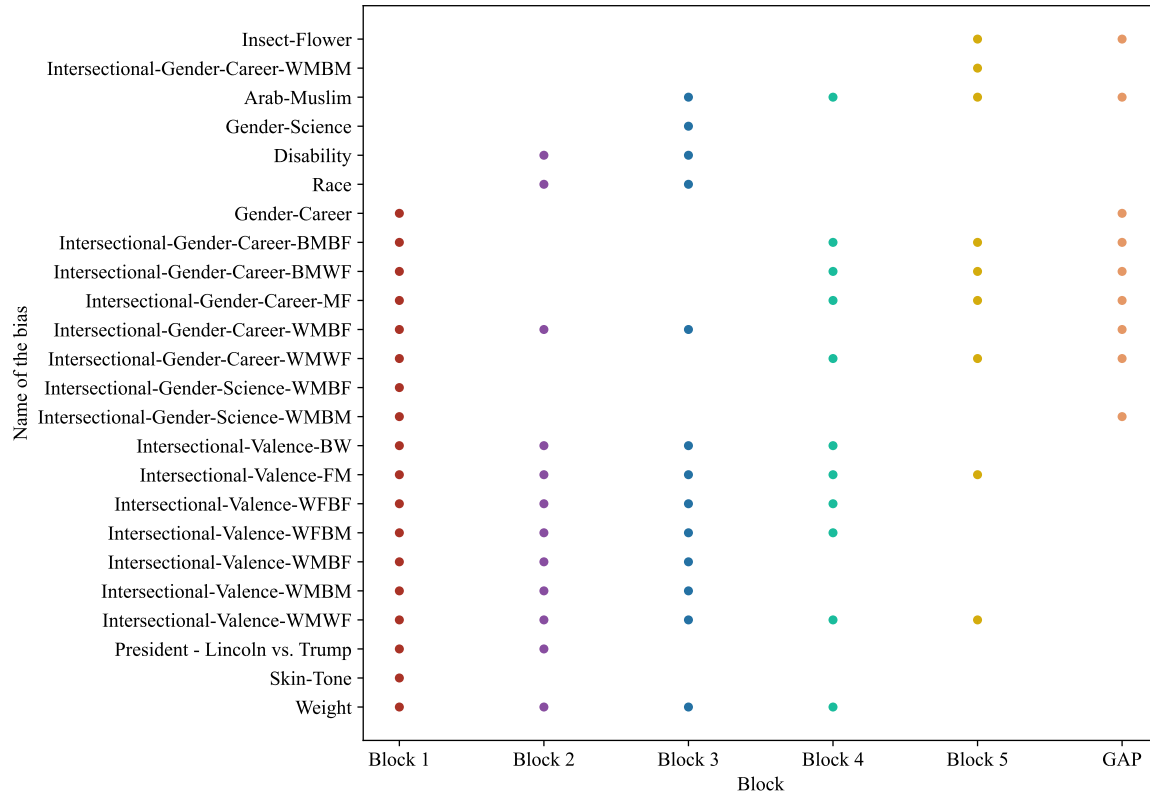


Figure 12. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of ODC model.

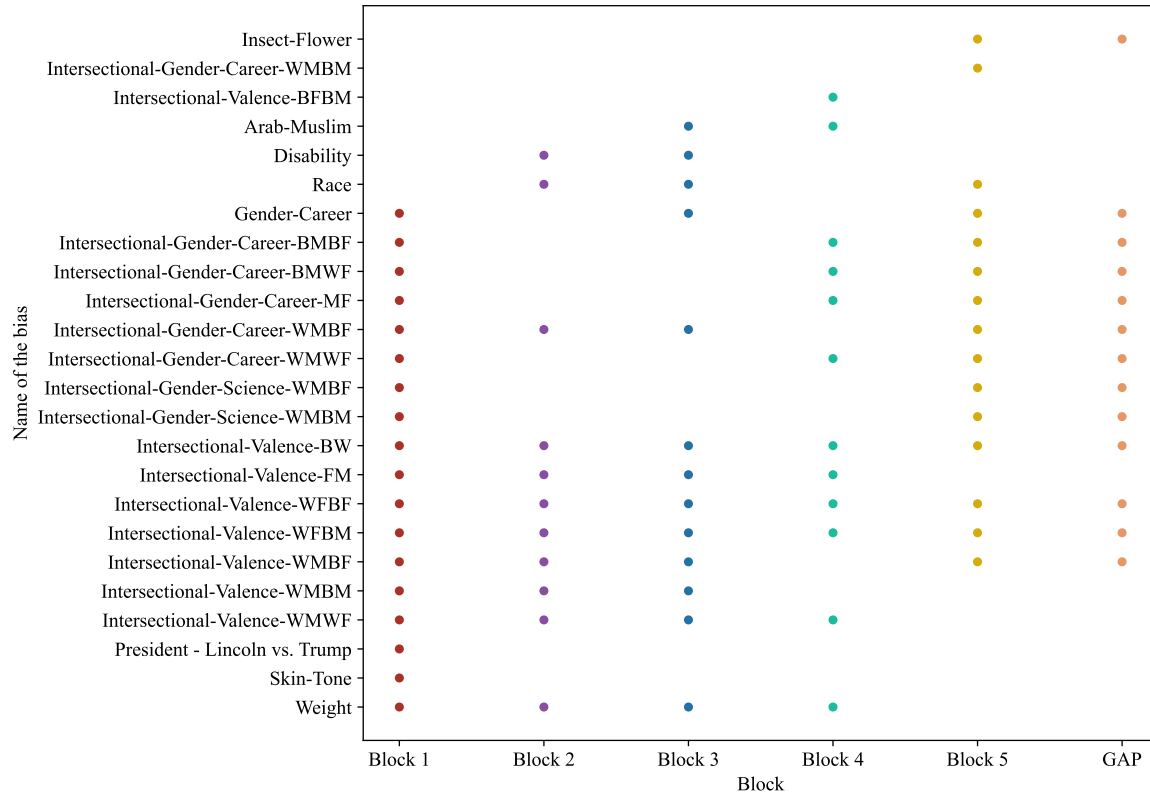


Figure 13. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of NPID model.

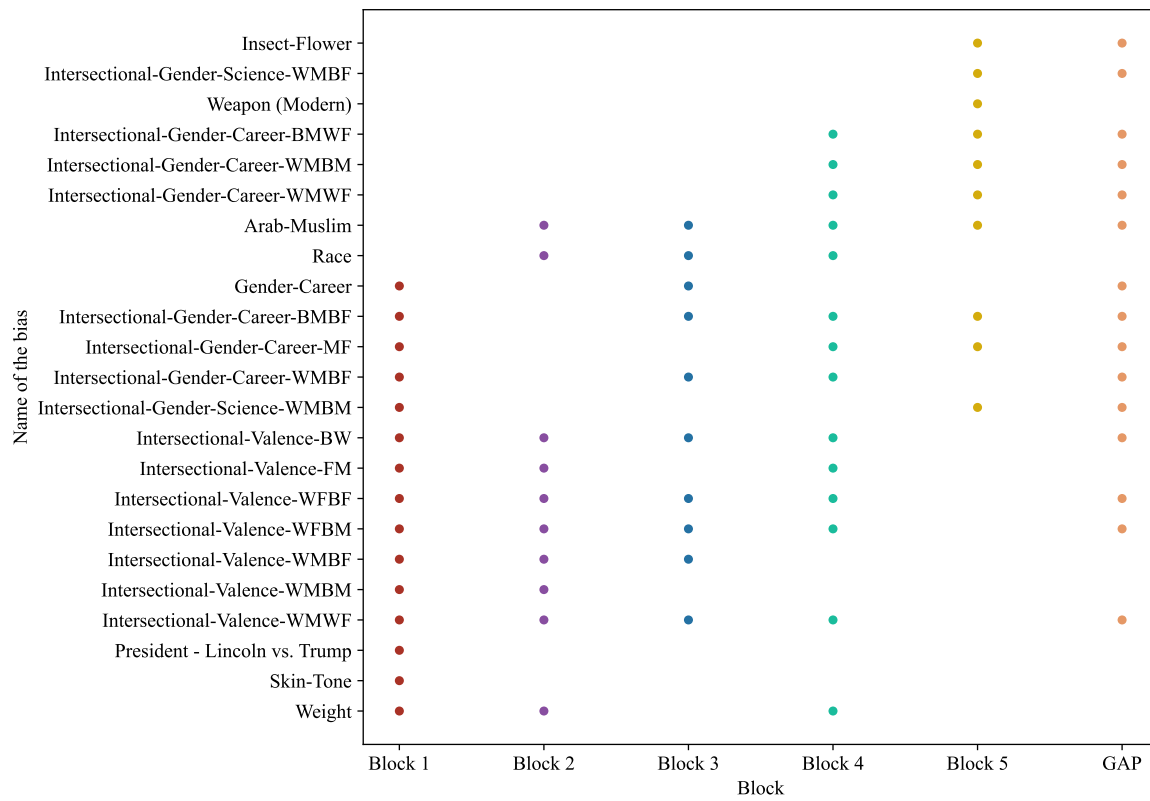


Figure 14. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of SimCLR model.

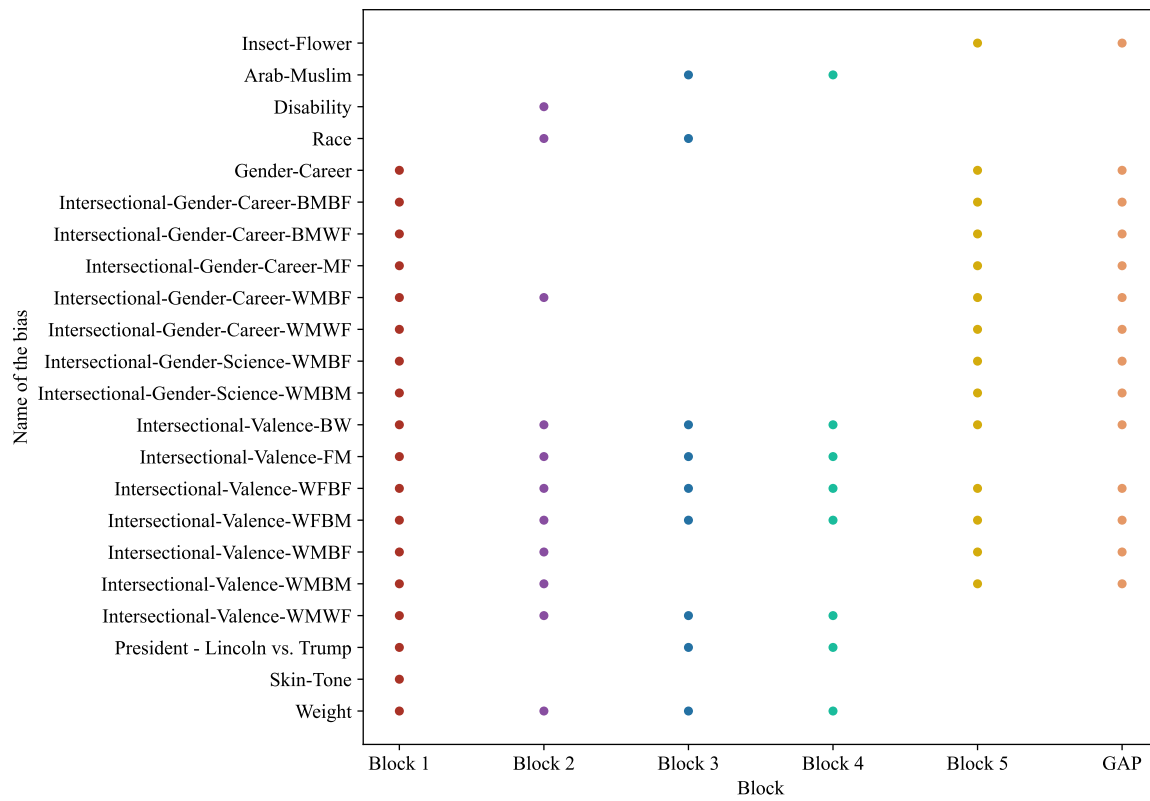


Figure 15. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of MoCo_v1 model.

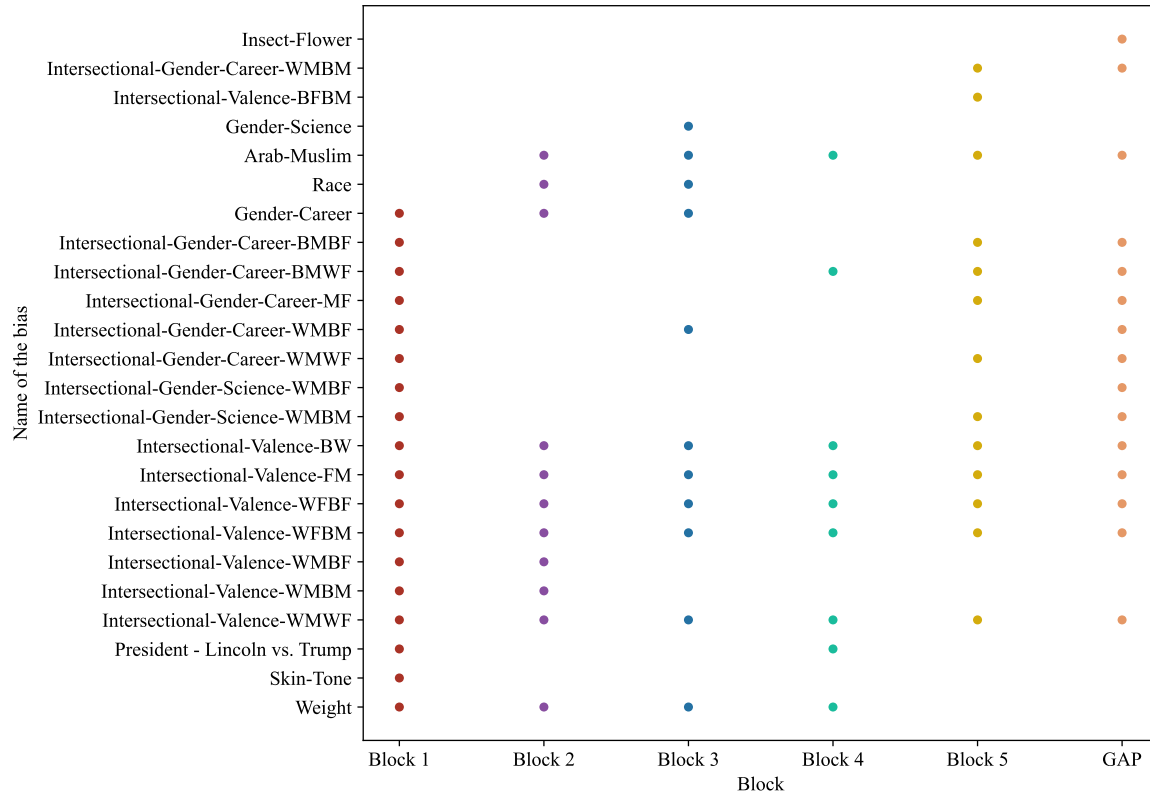


Figure 16. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of MoCo_v2 model.

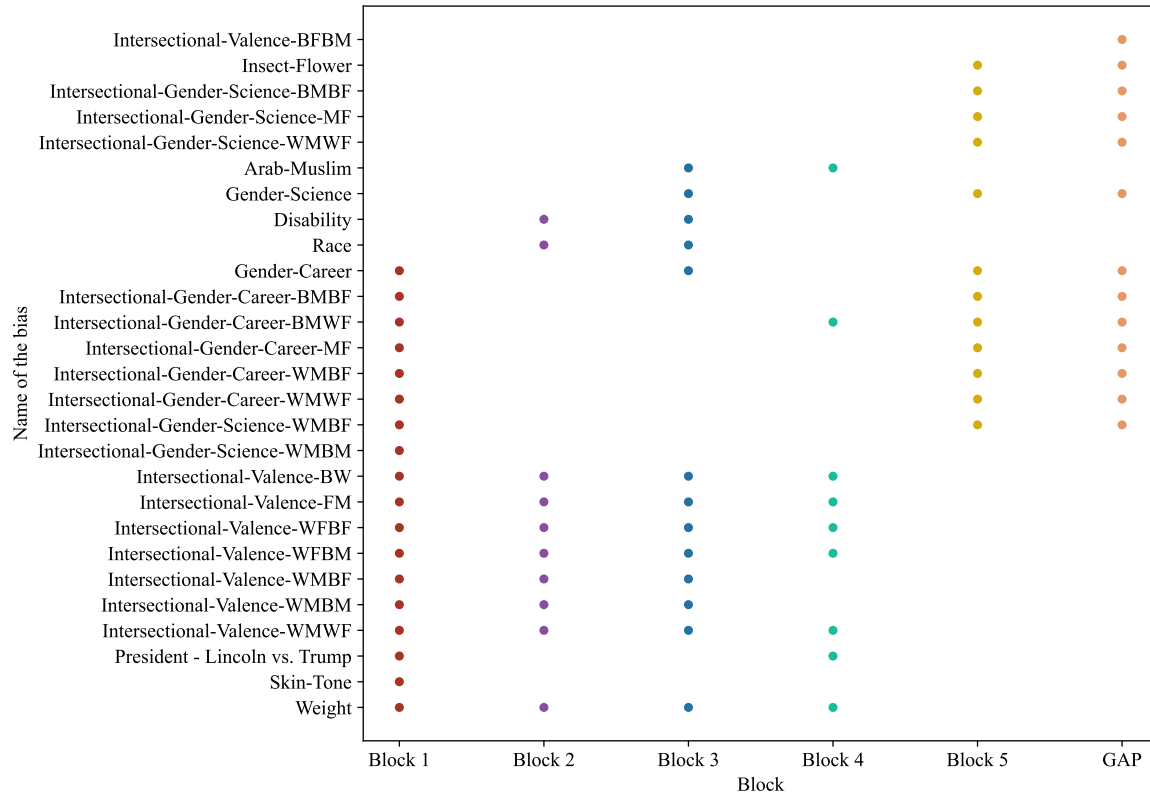


Figure 17. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the supervised model.

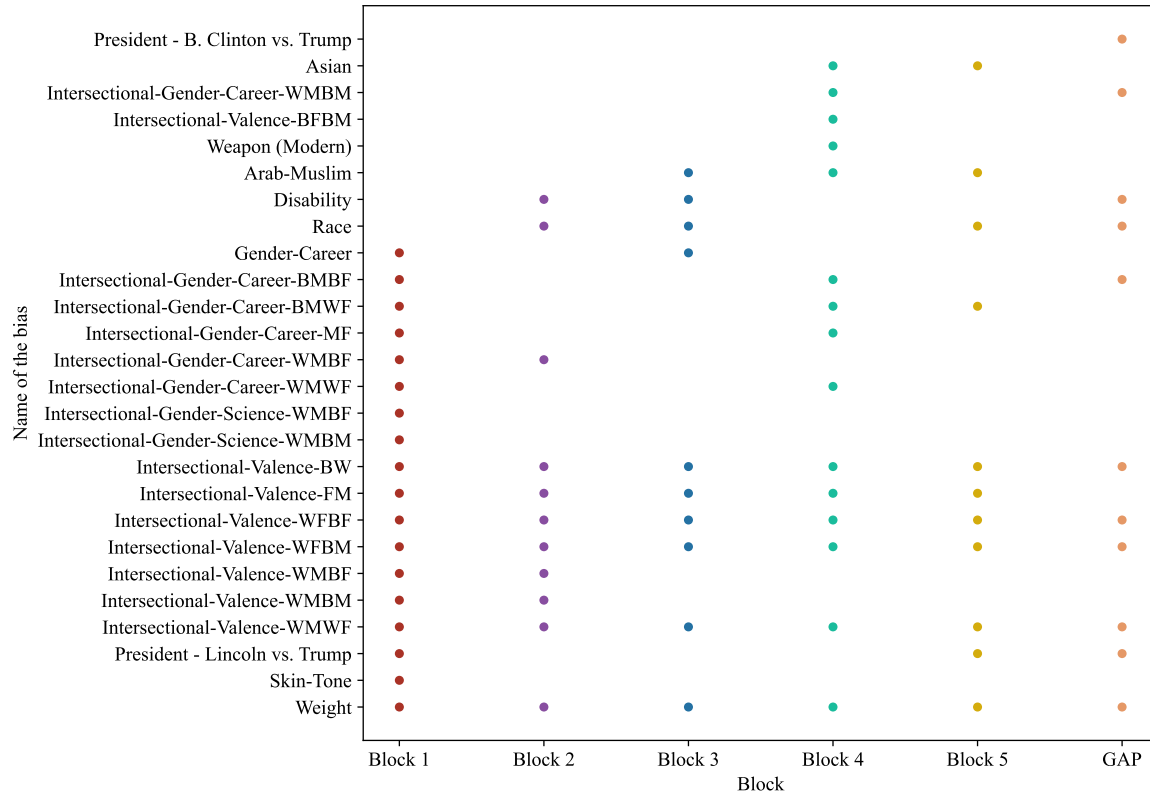


Figure 18. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the jigsaw model.

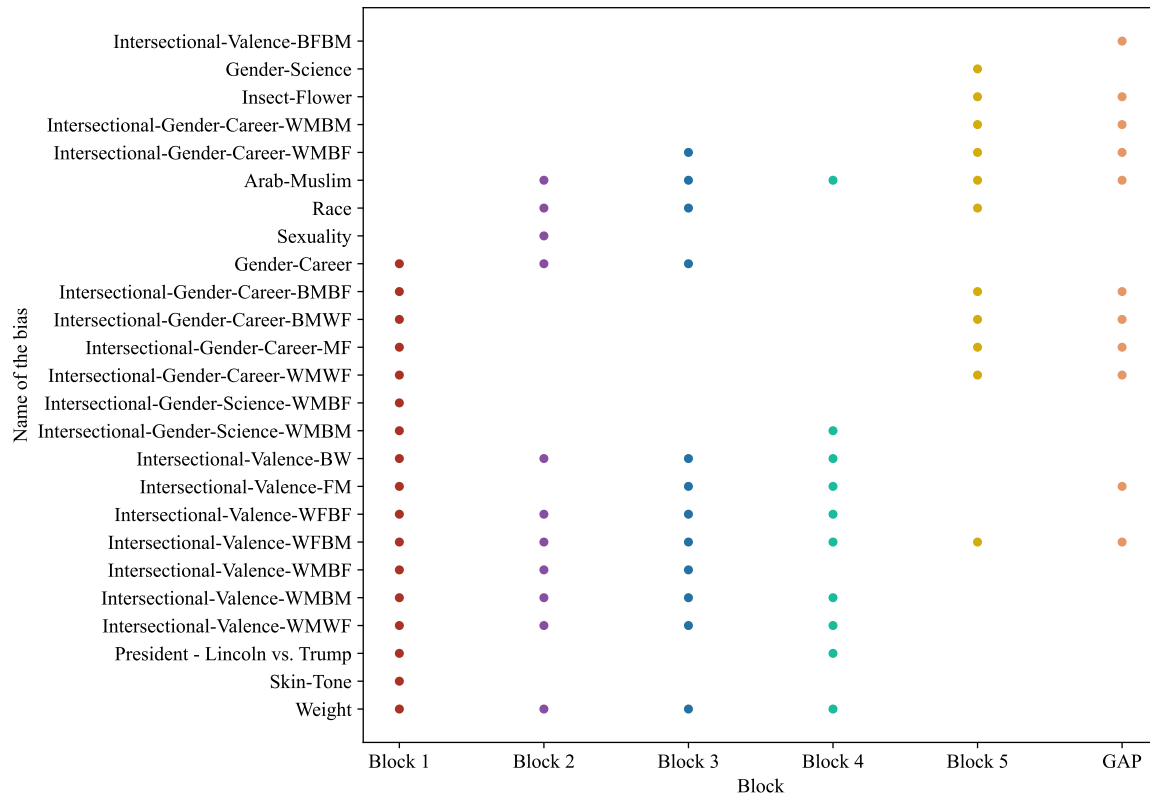


Figure 19. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the ClusterFit model.

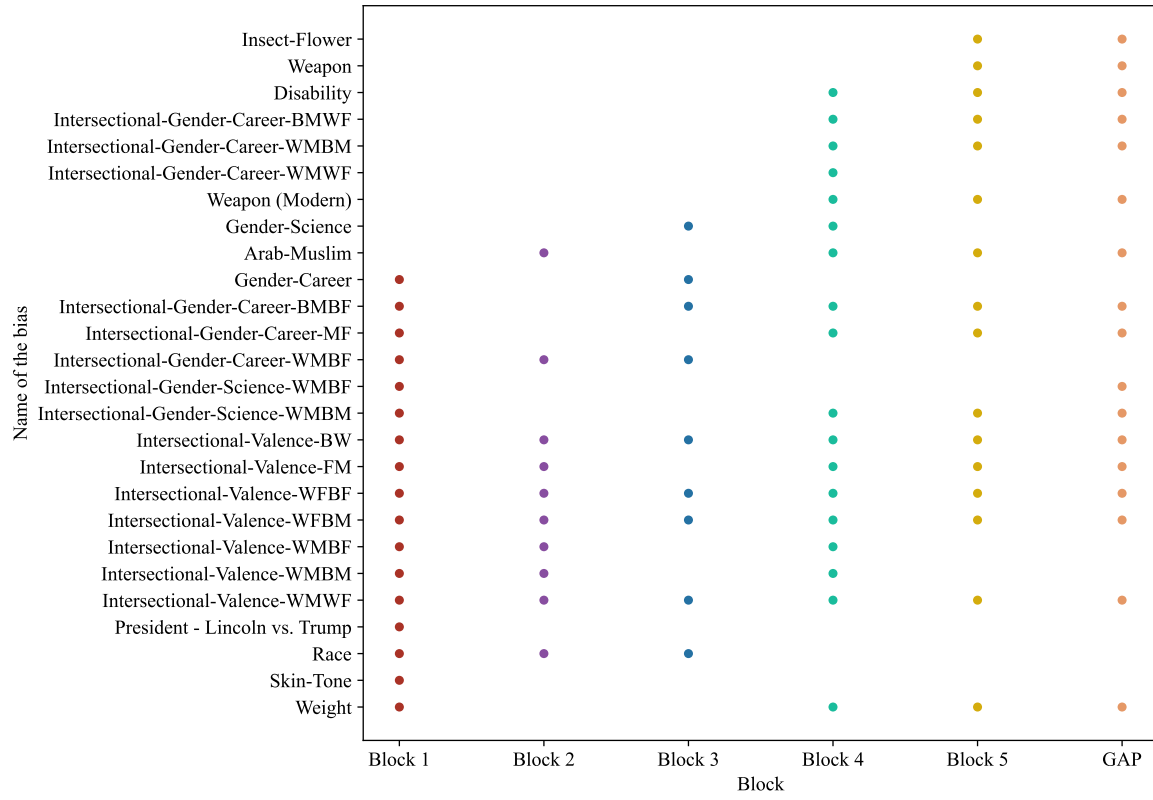


Figure 20. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the SwAV model.

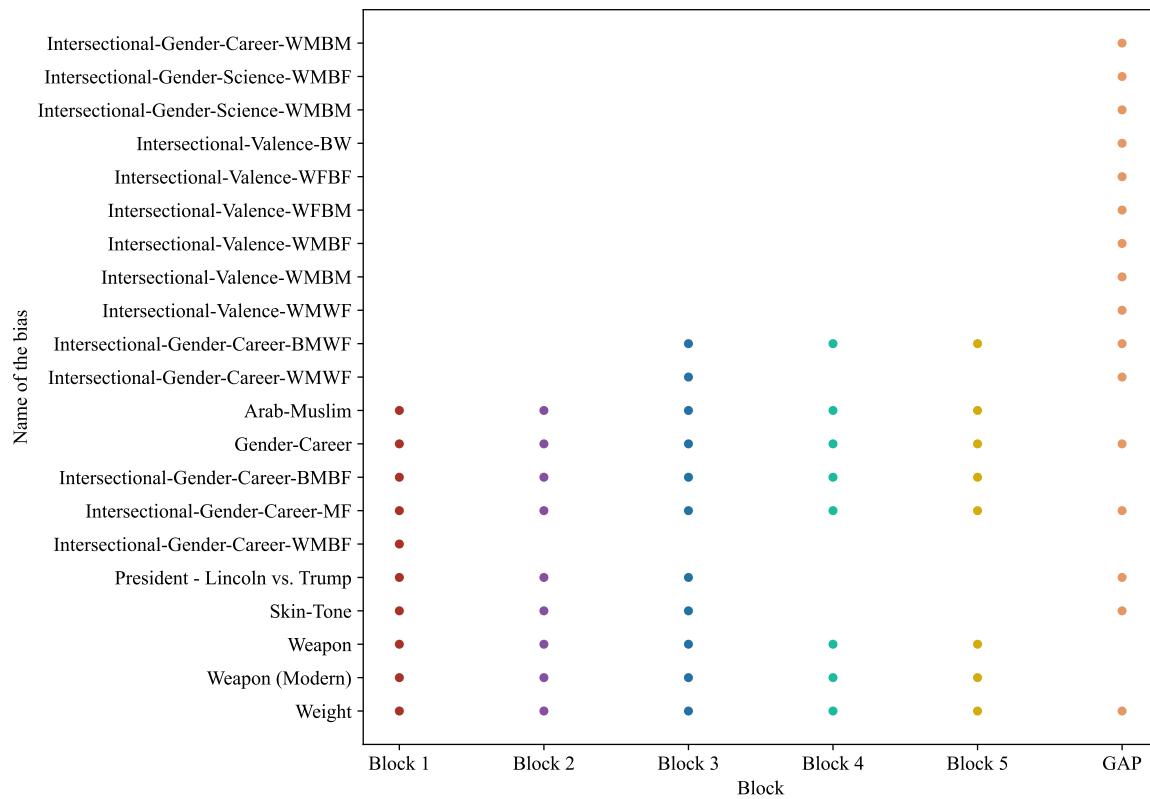


Figure 21. Presence of the statistically significant biases ($p_t < 0.1$) in the embeddings obtained from different layers of the random model.