

SUPPLEMENTARY TABLE 1. Page’s Trend Tests Across the Fixed Attire-Image Sequence.

Outcome	Male <i>z</i>	Male <i>p</i>	Female <i>z</i>	Female <i>p</i>
Trust	13.47	< 0.001	6.00	< 0.001
Choice likelihood	12.51	< 0.001	5.73	< 0.001
Technical competence	12.09	< 0.001	5.57	< 0.001
Professionalism	12.42	< 0.001	5.53	< 0.001
Communication comfort	12.05	< 0.001	4.53	< 0.001
Cleanliness/hygiene	10.99	< 0.001	5.87	< 0.001
Leadership	11.97	< 0.001	4.87	< 0.001
Experience	12.65	< 0.001	5.69	< 0.001
Uneventful surgery	11.93	< 0.001	5.21	< 0.001
Comfort when treated	11.56	< 0.001	5.75	< 0.001
Attractiveness	11.33	< 0.001	4.76	< 0.001
Global Perception Score	16.47	< 0.001	7.92	< 0.001
Context arm, presentation order (trust)	0.75	0.227	–	–

Page’s L test evaluates the a priori ordered alternative (casual < suit < suit with coat < scrubs < scrubs with white coat), which the Friedman test does not. *z* is the normal approximation; *p* is one-sided. The ordered hypothesis concerned the two pre-specified primary outcomes in the male-image primary analysis. For these two male-stream tests, *p* values were Bonferroni-adjusted for two comparisons (both adjusted *p* < 0.001). Page tests for the remaining male-stream outcomes and all Page tests in the secondary female-image stream are unadjusted and descriptive. For comparison, the corresponding test across the five surgical-context images in their presentation order is also shown as a sensitivity check and is likewise unadjusted; because those images differ from the attire images in content as well as position, this comparison cannot isolate an order effect from a content-domain effect.

SUPPLEMENTARY TABLE 2. Median (Interquartile Range) Ratings, Male-Image Stream (n = 189).

By attire	Casual	Suit	Suit + coat	Scrubs	Scrubs + coat
Trust	3 (2-4)	3 (3-4)	4 (3-5)	4 (4-5)	4 (4-5)
Choice likelihood	3 (2-4)	3 (3-5)	4 (3-5)	4 (4-5)	4 (4-5)
Technical competence	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Professionalism	3 (2-4)	4 (2-5)	4 (3-5)	4 (4-5)	4 (4-5)
Communication comfort	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Cleanliness/hygiene	3 (2-4)	4 (2-4)	4 (3-5)	4 (3-5)	5 (4-5)
Leadership	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Experience	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Uneventful surgery	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Comfort when treated	3 (2-4)	3 (2-4)	4 (3-5)	4 (3-5)	4 (4-5)
Attractiveness	3 (1-4)	3 (2-4)	4 (3-4)	4 (3-5)	4 (3-5)
By surgical context	Neutral	Open	Endo	Lap	Robotic
Trust	4 (3-5)	4 (4-5)	4 (4-5)	4 (3-5)	4 (4-5)
Choice likelihood	4 (3-5)	4 (4-5)	4 (4-5)	4 (3-5)	4 (4-5)
Technical competence	4 (3-5)	4 (4-5)	4 (3-5)	4 (3-5)	4 (3-5)
Professionalism	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)
Communication comfort	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)
Cleanliness/hygiene	4 (3-5)	4 (4-5)	4 (3-5)	4 (3-5)	5 (3-5)
Leadership	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)
Experience	4 (3-5)	4 (4-5)	4 (3-5)	4 (3-5)	4 (3-5)
Uneventful surgery	4 (3-5)	4 (4-5)	4 (3-5)	4 (3-5)	4 (3-5)
Comfort when treated	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)	4 (4-5)
Attractiveness	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)

SUPPLEMENTARY TABLE 3. Pairwise Comparisons Across Attire, Male-Image Stream (n = 189).

Contrast	Trust <i>p</i>	<i>r</i>	Choice <i>p</i>	<i>r</i>	GPS <i>p</i>	<i>r</i>
Casual vs. suit	< 0.001	+0.59	< 0.001	+0.60	< 0.001	+0.69
Casual vs. suit + coat	< 0.001	+0.87	< 0.001	+0.83	< 0.001	+0.88
Casual vs. scrubs	< 0.001	+0.90	< 0.001	+0.90	< 0.001	+0.91
Casual vs. scrubs + coat	< 0.001	+0.92	< 0.001	+0.90	< 0.001	+0.91
Suit vs. suit + coat	< 0.001	+0.88	< 0.001	+0.75	< 0.001	+0.86
Suit vs. scrubs	< 0.001	+0.86	< 0.001	+0.83	< 0.001	+0.85
Suit vs. scrubs + coat	< 0.001	+0.93	< 0.001	+0.87	< 0.001	+0.92
Suit + coat vs. scrubs	0.112	+0.32	0.016	+0.40	0.014	+0.34
Suit + coat vs. scrubs + coat	< 0.001	+0.72	< 0.001	+0.75	< 0.001	+0.73
Scrubs vs. scrubs + coat	< 0.001	+0.67	< 0.001	+0.58	< 0.001	+0.77

Wilcoxon signed-rank tests with Bonferroni correction across the ten pairwise contrasts, reported with the matched-pairs rank-biserial correlation *r* as a paired effect size (positive *r* indicates higher ratings for the second-listed category). GPS, exploratory Global Perception Score.

SUPPLEMENTARY TABLE 4. Internal Consistency and Inter-Item Correlations.

	Tru	Cho	Com	Pro	Cmf	Hyg	Lea	Exp	Unv	Trt
Trust	1.00	0.91	0.85	0.85	0.82	0.83	0.80	0.81	0.80	0.85
Choice likelihood	0.91	1.00	0.85	0.85	0.82	0.84	0.81	0.81	0.80	0.86
Technical competence	0.85	0.85	1.00	0.84	0.83	0.82	0.83	0.82	0.82	0.84
Professionalism	0.85	0.85	0.84	1.00	0.82	0.84	0.80	0.81	0.80	0.84
Communication comfort	0.82	0.82	0.83	0.82	1.00	0.81	0.80	0.81	0.81	0.83
Cleanliness/hygiene	0.83	0.84	0.82	0.84	0.81	1.00	0.78	0.79	0.78	0.83
Leadership	0.80	0.81	0.83	0.80	0.80	0.78	1.00	0.86	0.84	0.83
Experience	0.81	0.81	0.82	0.81	0.81	0.79	0.86	1.00	0.84	0.83
Uneventful surgery	0.80	0.80	0.82	0.80	0.81	0.78	0.84	0.84	1.00	0.83
Comfort when treated	0.85	0.86	0.84	0.84	0.83	0.83	0.83	0.83	0.83	1.00

Cronbach’s alpha computed separately within each image condition ranged from 0.972 to 0.989 (median 0.984); the value pooled across all conditions was 0.985. Because the within-condition and pooled values are almost identical, the high alpha does not arise from pooling across conditions but reflects genuine redundancy among the items. Values below are mean Spearman correlations averaged across the ten image conditions. Mean inter-item correlation 0.83 (range 0.78-0.91). Tru, trust; Cho, choice likelihood; Com, technical competence; Pro, professionalism; Cmf, communication comfort; Hyg, cleanliness/hygiene; Lea, leadership; Exp, experience; Unv, uneventful surgery; Trt, comfort when treated.

SUPPLEMENTARY TABLE 5. Physician Attire and Patient Perceptions, Female-Image Stream (n = 31).

Perception item	Casual	Suit	Suit + coat	Scrubs	Scrubs + coat	χ^2	<i>p</i> *	W	Page <i>z</i>
Trust	2.61 ± 1.28	3.13 ± 1.28	3.52 ± 1.12	3.77 ± 1.09	4.00 ± 1.10	51.9	< 0.001	0.419	6.0
Choice likelihood	2.74 ± 1.44	3.29 ± 1.51	3.84 ± 1.27	4.03 ± 1.20	4.10 ± 1.08	54.6	< 0.001	0.440	5.7
Technical competence	2.61 ± 1.33	3.23 ± 1.48	3.48 ± 1.03	3.81 ± 1.05	4.13 ± 1.18	45.2	< 0.001	0.364	5.6
Professionalism	2.68 ± 1.38	3.32 ± 1.47	3.71 ± 1.07	4.03 ± 1.08	4.13 ± 1.09	49.8	< 0.001	0.402	5.5
Communication comfort	2.81 ± 1.42	3.29 ± 1.47	3.90 ± 1.16	3.87 ± 1.15	4.03 ± 1.14	37.3	< 0.001	0.301	4.5
Cleanliness/hygiene	2.58 ± 1.29	3.32 ± 1.35	3.65 ± 0.95	3.97 ± 1.02	4.06 ± 1.09	51.9	< 0.001	0.418	5.9
Leadership	2.71 ± 1.37	3.29 ± 1.24	3.74 ± 1.18	3.87 ± 1.12	3.87 ± 1.12	40.2	< 0.001	0.324	4.9
Experience	2.61 ± 1.36	3.32 ± 1.40	3.65 ± 1.23	3.90 ± 1.14	4.13 ± 1.15	54.0	< 0.001	0.436	5.7
Uneventful surgery	2.61 ± 1.31	3.23 ± 1.33	3.61 ± 1.05	3.84 ± 1.13	3.90 ± 1.01	43.8	< 0.001	0.354	5.2
Comfort when treated	2.58 ± 1.43	3.19 ± 1.40	3.81 ± 1.08	3.90 ± 1.16	4.03 ± 1.17	53.3	< 0.001	0.430	5.7
Attractiveness	2.61 ± 1.23	3.23 ± 1.33	3.42 ± 1.06	3.61 ± 1.12	3.74 ± 1.12	36.6	< 0.001	0.296	4.8
Would proceed, % yes	51.6	58.1	83.9	87.1	93.5	Q = 31.3	< 0.001	–	–

Secondary analysis. Values are mean ± standard deviation. This stream is reported separately and is not compared with the male-image stream, because image gender was determined by participant preference rather than randomized. **p*, Bonferroni-adjusted across the eleven outcomes. Bottom row: binary treatment acceptance, Cochran Q test.

SUPPLEMENTARY TABLE 6. Displayed Surgical Context and Patient Perceptions, Female-Image Stream (n = 31).

Perception item	Neutral	Open	Endo	Lap	Robotic	χ^2	<i>p</i> *	W
Trust	4.06 ± 1.15	4.00 ± 1.26	3.94 ± 1.41	3.77 ± 1.48	4.06 ± 1.41	2.9	1.000	0.023
Choice likelihood	4.00 ± 1.21	4.06 ± 1.24	4.00 ± 1.39	3.74 ± 1.57	4.16 ± 1.29	4.4	1.000	0.035
Technical competence	3.94 ± 1.18	4.10 ± 1.19	4.10 ± 1.40	3.84 ± 1.59	4.19 ± 1.25	4.0	1.000	0.032
Professionalism	4.06 ± 1.29	4.19 ± 1.14	3.97 ± 1.56	3.90 ± 1.51	4.13 ± 1.23	2.6	1.000	0.021
Communication comfort	4.00 ± 1.26	4.29 ± 1.10	4.10 ± 1.35	3.84 ± 1.57	4.13 ± 1.23	6.6	1.000	0.053
Cleanliness/hygiene	4.10 ± 1.14	4.13 ± 1.26	3.81 ± 1.54	3.81 ± 1.62	4.10 ± 1.25	4.4	1.000	0.036
Leadership	4.03 ± 1.17	4.10 ± 1.30	4.13 ± 1.31	3.97 ± 1.40	4.06 ± 1.29	0.4	1.000	0.003
Experience	4.10 ± 1.19	4.13 ± 1.31	3.97 ± 1.49	3.84 ± 1.55	4.13 ± 1.20	3.3	1.000	0.026
Uneventful surgery	3.97 ± 1.14	4.00 ± 1.24	4.00 ± 1.32	3.84 ± 1.49	4.23 ± 1.18	5.9	1.000	0.048
Comfort when treated	3.90 ± 1.19	4.00 ± 1.37	4.10 ± 1.35	3.87 ± 1.52	4.10 ± 1.19	4.8	1.000	0.038
Attractiveness	3.90 ± 1.33	3.90 ± 1.27	3.81 ± 1.47	3.65 ± 1.62	4.03 ± 1.20	2.3	1.000	0.018
Would proceed, % yes	93.5	87.1	90.3	77.4	93.5	Q = 7.2	0.127	–

**p*, Bonferroni-adjusted across the eleven outcomes. No comparison reached significance.

SUPPLEMENTARY TABLE 7. Sensitivity Analyses for the Model of Treatment Acceptance.

Variable	GEE exchangeable (primary)	GEE independence	Firth penalized	Random intercept
Position 2: suit (vs. position 1: casual)	1.90 (1.44-2.50)	1.88 (1.44-2.46)	1.86 (1.19-2.89)	3.26 (1.70-6.24)
Position 3: suit + coat (vs. position 1: casual)	14.94 (7.56-29.52)	14.58 (7.57-28.08)	13.48 (6.74-26.99)	83.74 (27.46-255.44)
Position 4: scrubs (vs. position 1: casual)	21.84 (9.49-50.25)	21.29 (9.47-47.87)	19.26 (8.74-42.41)	139.39 (40.65-478.03)
Position 5: scrubs + coat (vs. position 1: casual)	39.10 (13.11-116.66)	38.08 (12.86-112.71)	32.73 (12.33-86.92)	293.69 (68.45-1260.08)
Age (per 10 years)	0.92 (0.74-1.14)	0.90 (0.74-1.10)	0.90 (0.77-1.05)	0.84 (0.58-1.21)
Participant sex: female (vs. male)	2.08 (0.82-5.27)	1.23 (0.46-3.27)	1.22 (0.68-2.16)	1.71 (0.44-6.61)
Secondary education (vs. primary)	1.72 (0.70-4.24)	1.76 (0.77-4.02)	1.74 (0.92-3.27)	2.62 (0.56-12.28)
University (vs. primary)	2.75 (0.86-8.78)	2.79 (0.95-8.14)	2.72 (1.24-5.97)	6.66 (0.99-44.55)
Postgraduate (vs. primary)	4.07 (0.99-16.67)	3.05 (0.83-11.18)	2.94 (1.10-7.82)	7.01 (0.68-72.63)
Prior urological surgery (vs. none)	1.81 (0.92-3.53)	1.69 (0.89-3.22)	1.68 (1.06-2.65)	2.70 (0.90-8.12)
Medium perceived income (vs. low)	1.54 (0.75-3.15)	1.35 (0.70-2.61)	1.35 (0.80-2.29)	1.90 (0.54-6.76)
High perceived income (vs. low)	0.61 (0.20-1.83)	0.55 (0.19-1.55)	0.56 (0.27-1.15)	0.41 (0.07-2.35)

All models use the same covariates and the same analytic sample (915 observations from 183 participants). The primary model is the GEE with an exchangeable working correlation (Table 5). Firth penalized likelihood was fitted using a custom implementation (Python/NumPy/SciPy) without adjustment for within-participant clustering; confidence intervals were calculated using Wald standard errors obtained from the inverse penalized information matrix. Because each participant contributed five attire responses, this model does not itself provide valid standard errors for repeated measurements. The Firth model is presented only as a limited sensitivity check for sparse-data and near-separation behaviour; its similarity in direction and ordering to the GEE estimates should not be interpreted as a clustering-adjusted validation of the GEE model. The random-intercept model (statsmodels-based adaptive Gauss-Hermite quadrature) estimates subject-specific (conditional) odds ratios, which are expected to be larger in magnitude than the population-averaged (marginal) odds ratios of the GEE models when between-participant variation is substantial (random-intercept SD: 2.45); conditional and marginal odds ratios are therefore not directly comparable in size, only in direction and ordering. Across all four models, the direction and large magnitude of the contrasts across the fixed attire-image sequence were broadly consistent, but only the GEE and random-intercept models appropriately account for repeated measurements within participants. Values are odds ratios (95% confidence intervals). SD, standard deviation.

CONTINUED SUPPLEMENTARY TABLE 7. Adjusted probabilities and risk differences (primary GEE model)

Attire	Casual	Suit	Suit + coat	Scrubs	Scrubs + coat
Adjusted acceptance, %	57.6	71.0	94.6	96.2	97.8
Risk difference vs. casual, pp	–	+13.4	+37.0	+38.7	+40.3

pp, percentage points.