Supplementary Information

RAISE: A self-driving laboratory for interfacial property formulation discovery

Mohammad Nazeri [1†], Sheldon Mei [2†], Jeffrey Watchorn [3†], Alex Zhang [4], Erin Ng [2], Tao Wen [4], Abhijoy Mandal [5], Kevin Golovin [4], Alán Aspuru-Guzik [3,5,6,7], Frank Gu [1,2,3*]

Affiliations
[1] Institute of Biomedical Engineering, University of Toronto, 164 College St, Toronto, ON M5S 3G9
[2] Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College St, Toronto, ON M5S 3E5
[3] Acceleration Consortium, 700 University Avenue, Toronto, ON M5G 1X6
[4] Department of Mechanical & Industrial Engineering, University of Toronto, 5 King's College Rd, Toronto, ON M5S 3G8
[5] Department of Computer Science, University of Toronto, Toronto, Ontario, M5S 3H6, Canada
[6] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada
[7] Vector Institute for Artificial Intelligence, 661 University Ave. Suite 710, ON M5G 1M1, 15 Toronto, Canada

* Corresponding author:
Frank Gu

Corresponding author's email address
f.gu@utoronto.ca

To guide the Bayesian optimization toward user-defined targets, the RAISE interface allows objective values to be transformed into normalized desirability scores using three selectable transformation functions: linear, triangular, and bell-shaped transformations. These functions convert measured experimental responses into normalized values within the range [0,1], where higher values correspond to better agreement with the desired objective.

The linear transformation assigns desirability values that increase or decrease steadily across a specified range. Values closer to the preferred direction (higher or lower depending on the objective) receive higher desirability scores. This transformation is typically used when the goal is to monotonically maximize or minimize a parameter. The triangular transformation defines a single optimal target value. The desirability is highest at the target and decreases linearly as the measured value deviates from it. This transformation is useful when the objective is to match a specific value rather than simply increase or decrease a parameter. The bell-shaped transformation applies a smooth weighting centered around the target value. Values closest to the target receive the highest desirability, while deviations are penalized gradually. Compared to the triangular transformation, this approach provides a smoother reduction in desirability as the measured value moves away from the target.

Graphical user interface used to configure RAISE campaigns, shown in Figure S1. Users define reagent types, concentration ranges, and optimization objectives such as target contact angle and total reagent usage. These inputs are passed to the Bayesian Optimization agent, which iteratively proposes new formulations for testing.

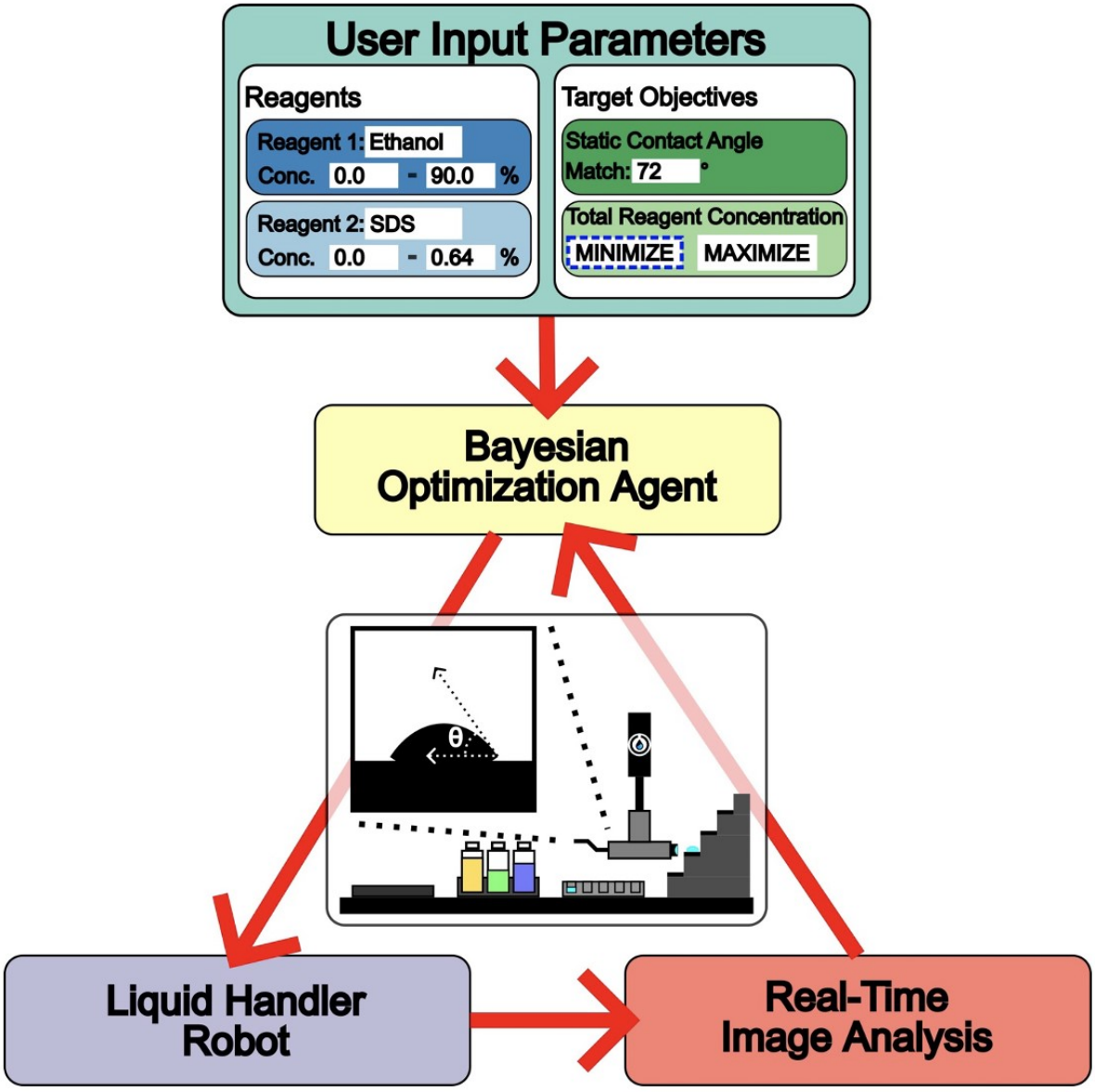

Figure S1. User interface and workflow for Bayesian optimization-driven formulation experiments in the RAISE platform.

To assess how consistent the contact angle measurements were, the standard deviations from both RAISE and the manual goniometer were compared across three different surfaces. An F-test was used to determine if one method was actually more reliable than the other, or if any differences observed were just random (Table S1). The F-test compares the variances (the standard deviations (SD) squared) between the two methods, and the p-values indicate how likely it is that these differences would occur just by chance. RAISE showed more consistent results than the manual goniometer across all surfaces tested. This was especially evident with PTFE, where RAISE's measurements were significantly more reliable (p-value = 0.0106). This improvement occurs

because RAISE automates the droplet placement and image capture, which eliminates much of the human error that can occur during manual operation.

Table S1. Comparison of measurement variability between RAISE and manual goniometer.

Substrate	RAISE SD (°)	Goniometer SD (°)	F-statistic	p-value	% Reduction in SD
PDMS	0.61	0.77	1.593	0.2243	20.8
PS	1.39	1.96	1.988	0.2155	29.1
PTFE	1.69	4.43	6.781	0.0106	61.9

To evaluate the consistency and stability of our contact angle measurement system, 3 µL water droplets were deposited onto PDMS substrates positioned at five different stage levels (A to E). Each level was tested across nine sample positions to account for variations across a single substrate. The measured contact angles (°) are reported in Table S2. This assessment was conducted to estimate the system's standard deviation (0.805) and ensure reliability across different measurement zones. The average contact angles observed in this dataset are slightly lower than those shown in the benchmarking results (Figure 5 of the main manuscript). This minor discrepancy is mainly due to the gradual degradation of this specific PDMS surface over time, which affects its hydrophobicity.

Table S2. Contact angle (°C) of 3 µL water droplets on PDMS at five stage levels. Standard deviation: 0.805°.

	Contact Angle (°C)								
Stage Level	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9
A	100.947	100.211	100.876	100.601	99.193	100.229	99.868	98.127	99.54
B	99.918	99.096	99.489	99.25	100.196	100.275	99.886	99.87	100.277
C	99.902	99.474	99.423	99.643	100.016	99.951	100.244	99.624	100.146
D	101.359	100.421	100.146	100.275	100.194	100.407	100.713	100.292	100.277
E	101.995	101.77	101.601	100.485	100.997	101.502	100.887	101.341	102.034
Total Average	**100.288**								
Total Standard Deviation	**0.805**								

Additional benchmarking experiments were conducted to test how varying vertical z-offsets affected contact angle measurements on the stage. Overall, capturing images of water droplets on PDMS substrate at current imaging height (zero offset) showed the closest number with the target contact angle measured by a goniometer (103°) with low standard deviation. Small deviations from this height, either by slightly raising or lowering the camera by a few tenths of a millimeter, resulted in contact angle values that were consistently farther from the target. The additional benchmarking experiment data has been shown as below in the Supplementary Information section (Figure S2).

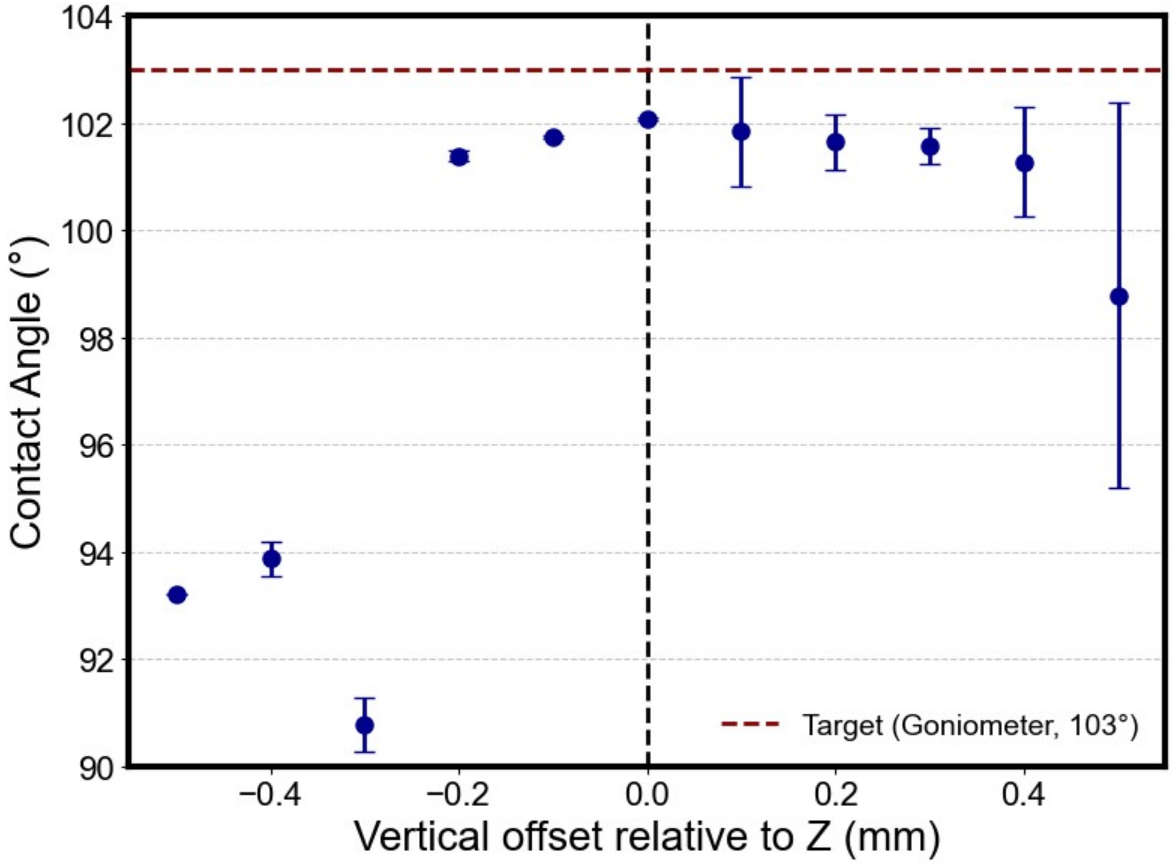

Figure S2. Effect of vertical camera offset on measured static contact angle. Measurements at the selected imaging height show the closest agreement with the goniometer reference.

To better demonstrate the platform's performance relative to the target contact angle across different experimental iterations, both the water-ethanol campaign (Figure 6 (a)) and the surfactant combination plot (Figure 8 (d)) are shown in Figure S3, where the y-axis represents the distance from the desired target contact angle.

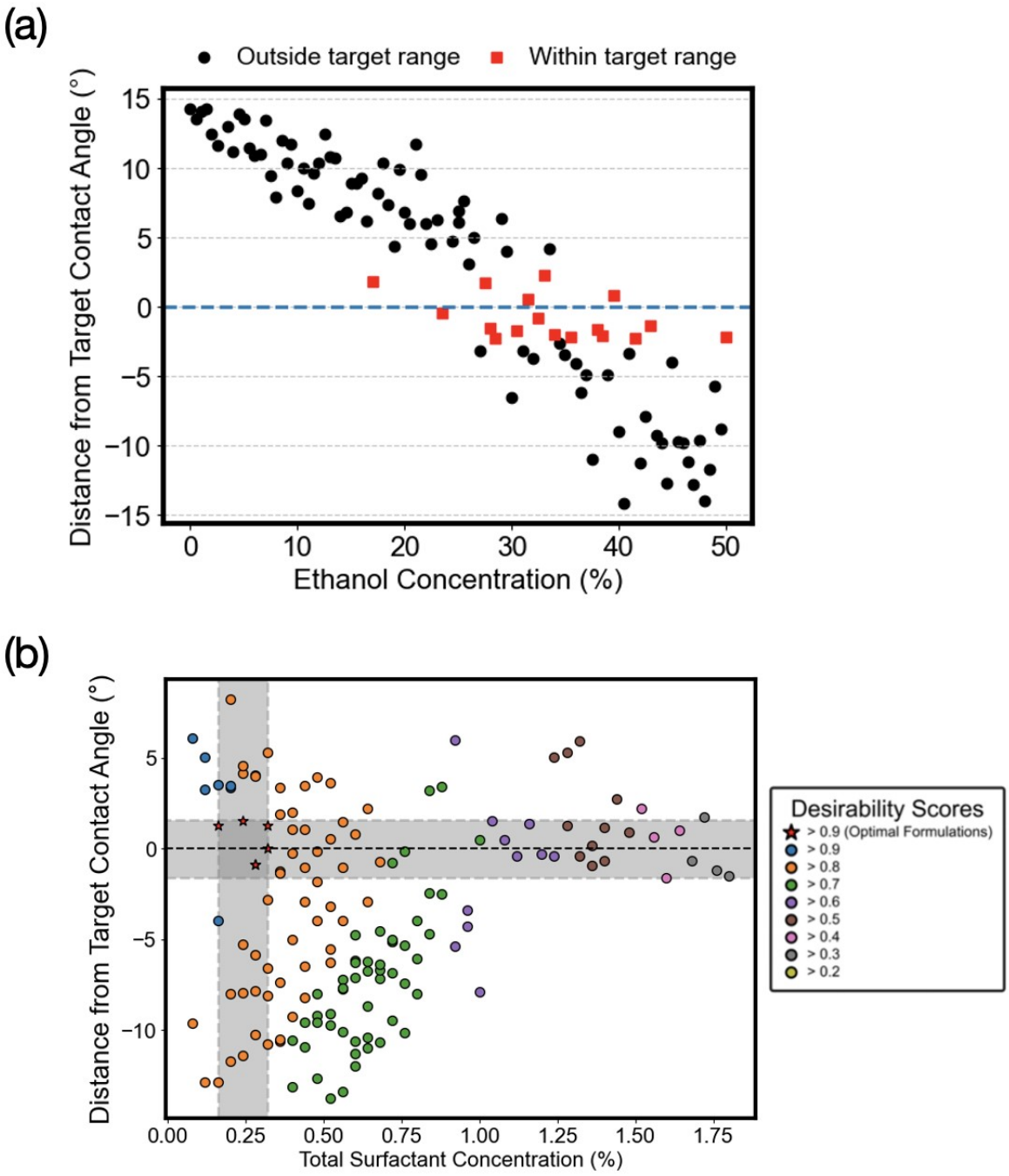

Figure S3. (a) Bayesian optimization driven campaign for water-ethanol mixture with a target of 85-90°. The y-axis is showing the distance of all iteration from target contact angle. (b) Combination of Tween 20 and SDS 94% campaign with showing the distance from target contact angle in the campaign (72°).

The RAISE user interface allows selection of linear, triangular, or bell-shaped transformations for target normalization. The linear transformation assigns desirability values that increase or decrease monotonically within specified bounds. The triangular transformation defines a maximum desirability at the target value, which decreases linearly toward defined lower and upper bounds. The bell transformation applies a Gaussian-like weighting centered at the target, enabling smooth penalization of deviations.

We conducted simulations with varying weight ratios between matching the contact angle and minimizing total surfactant usage to investigate the effects of prioritizing different design goals on optimization results. Testing whether prioritizing one goal over the other would affect the algorithm's capacity to identify the optimal formulations was the aim. The weights are presented in the format [target contact angle, minimal surfactant usage]. As shown in Figure S4, when both objectives are equally weighted ([1, 1]), the algorithm efficiently discovers formulations that balance performance and surfactant reduction. Increasing the weight on surfactant minimization (e.g., [1, 3] or [1, 5]) leads to leaner formulations but may sacrifice some accuracy in contact angle. On the other hand, more accurate wettability is achieved by giving priority to contact angle matching (e.g., [3, 1] or [5, 1]), but with a higher surfactant usage.

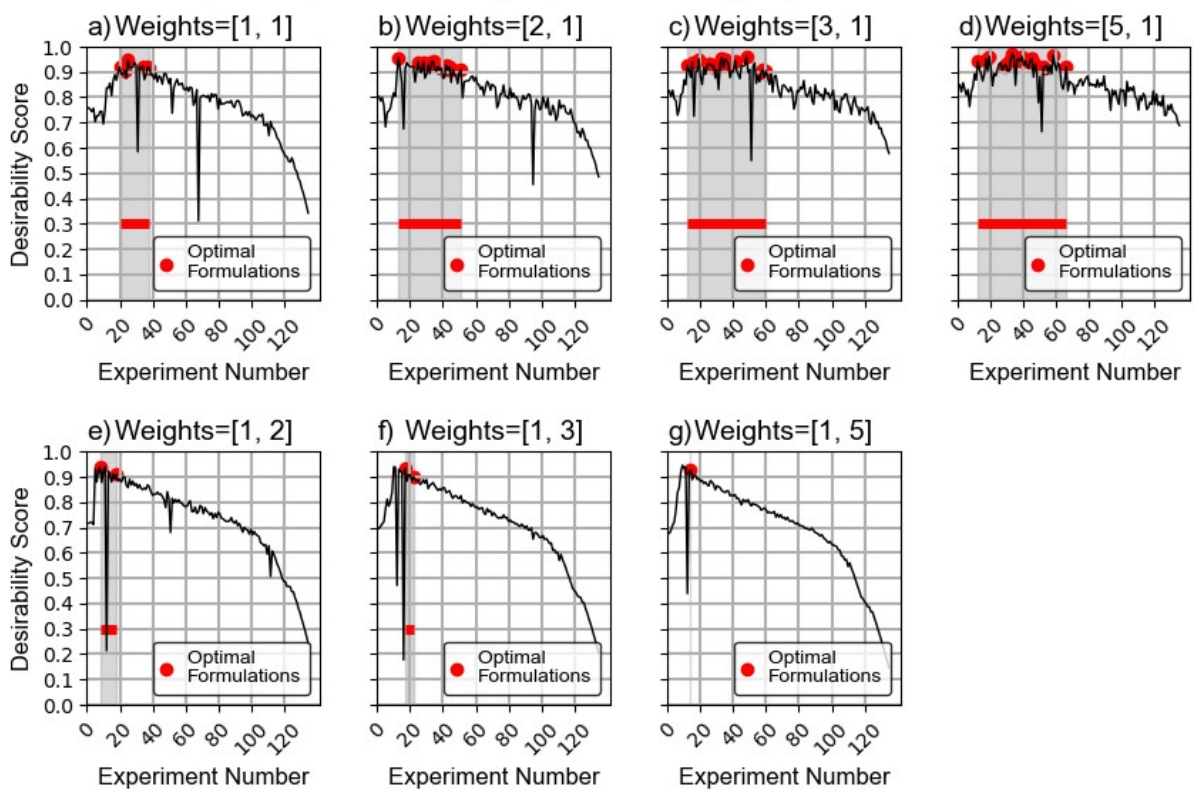

Figure S4. Impact of different weighting ratios between contact angle matching and surfactant minimization on BO performance. Giving equal ratio to both goals lead to balanced results, while shifting priorities alters the composition of optimal formulations, and the time necessary to identify them.

To understand how a strict contact angle exploration range affects the optimization process, we performed simulations using different exploration ranges around the target contact angle of 72°. These bounds ranged from very wide (±50°) to very narrow (±0.805°), which corresponds to two standard deviations based on experimental measurements. The goal was to explore how the precision of the design exploration range influences the ability of the optimization algorithm to find desirable formulations. As shown in the Figure S5, when the contact angle constraint is loose, the algorithm identifies high-desirability formulations more frequently and earlier in the campaign. In contrast, tighter bounds make it harder to meet the contact angle requirement, resulting in fewer qualifying formulations and longer discovery times.

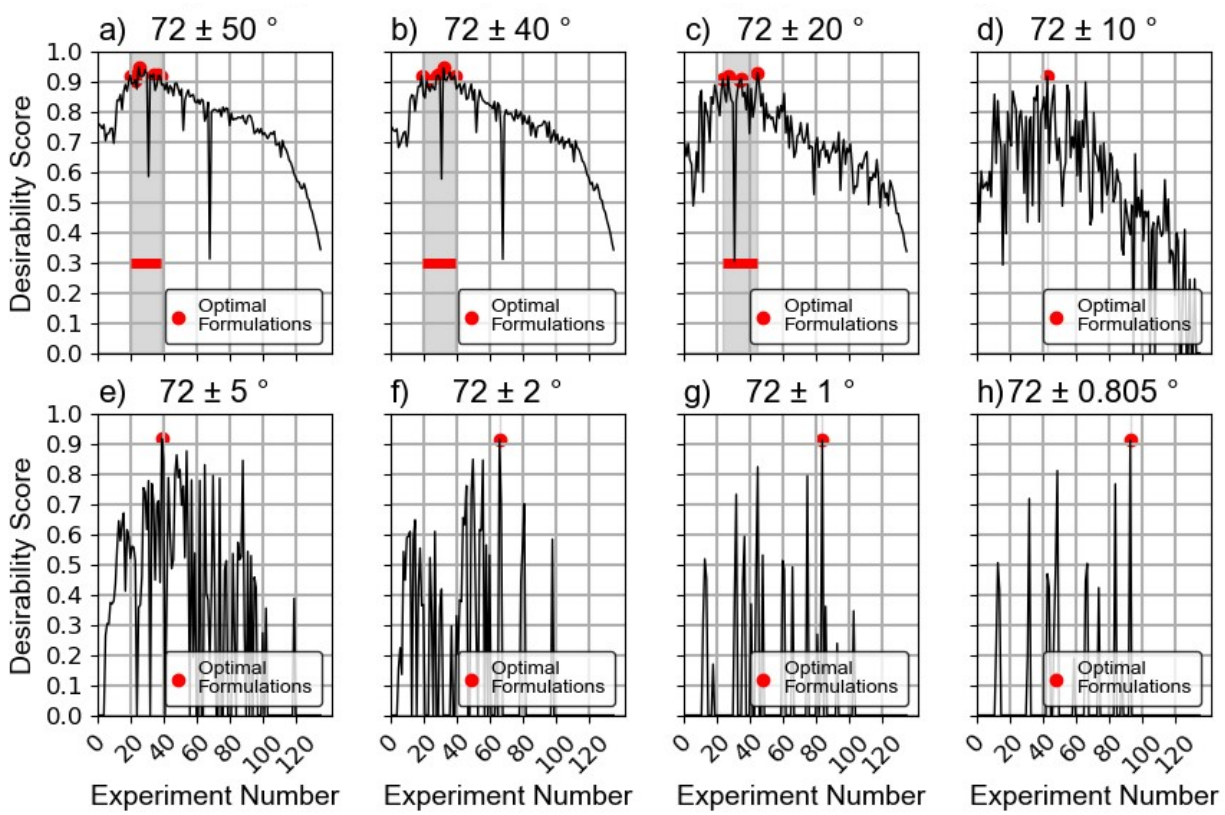

Figure S5. Effect of changing contact angle bounds on multi-objective BO performance. Stricter limits make it harder to find good formulations quickly, while wider ranges help the algorithm find solutions faster.

To investigate the effect of initial formulation selection, a simulated test was performed by varying the random starting formulation for both single- and multi-objective optimization. As show in Figure S6, each line shows how the desirability score changed over time for a different random starting point. The results demonstrate that multi-objective optimization was able to find optimal formulations (those with high desirability that also fall within the target contact angle rang) earlier and more consistently across all tests. This suggests that the multi-objective method is more efficient and less sensitive to how the experiment begins, making it more reliable in practice.

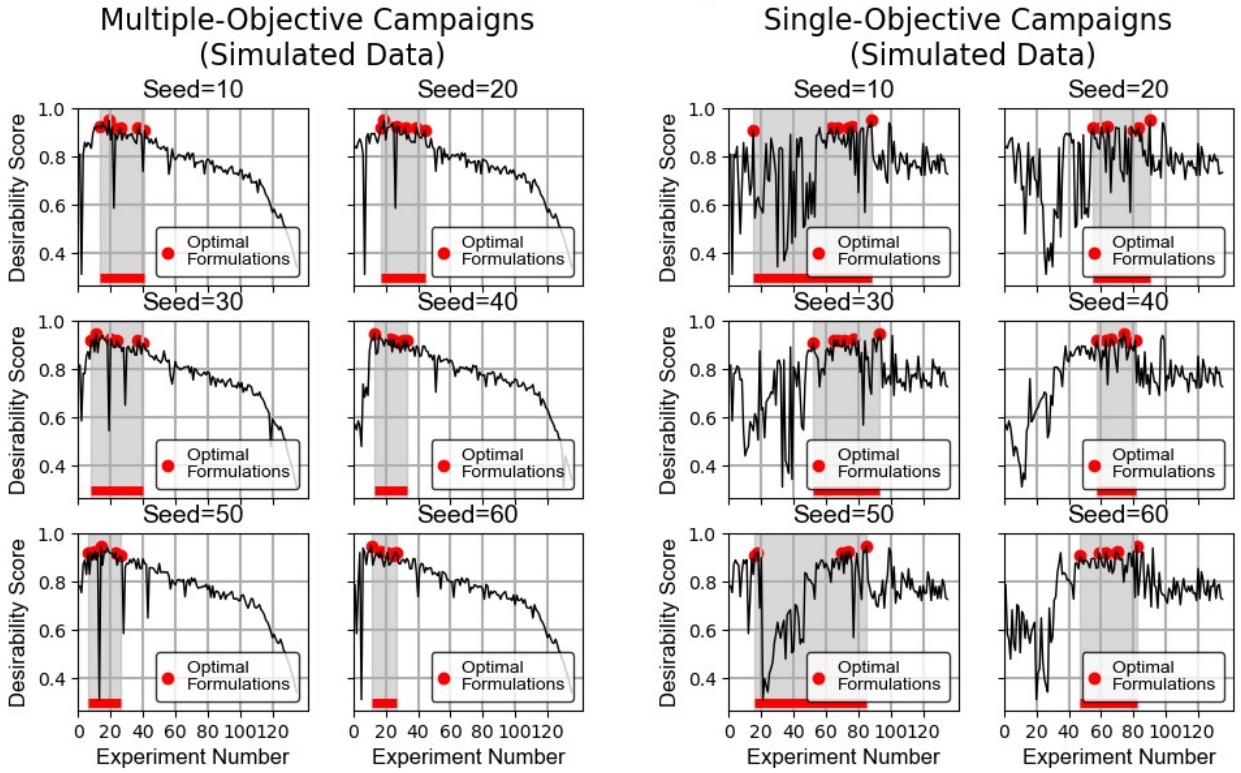

Figure S6. Comparison of multi-objective and single-objective optimization using different random starting formulations. Multi-objective optimization finds optimal formulations earlier and more consistently, no matter where it starts. Random seeds were set at the beginning of experiments to initialize the random number generator and to ensure reproducibility in the order of recommended experiments.

To visualize the formulation landscape explored by the BO campaign, we generated contour plots showing the variation of key target metrics-static contact angle, total surfactant concentration, and overall desirability-across the tested formulation space (Figure S7). Each plot maps SDS 94% and Tween 20 concentrations on the vertical and horizontal axes, respectively. Figure S7 (a) shows how the contact angle varies across the formulation space, with a broad plateau region around 72°, matching the desired wettability profile of high-purity SDS (99%). Figure S7 (b) illustrates the total surfactant concentration, which increases linearly with the combined component inputs, and serves as a penalty metric in the desirability function. Figure S7 (c) presents the computed desirability score, which combines both target parameters into a single optimization objective. Notably, the optimal formulations (highlighted in red) cluster within a region of high desirability, low surfactant content, and contact angles near the target.

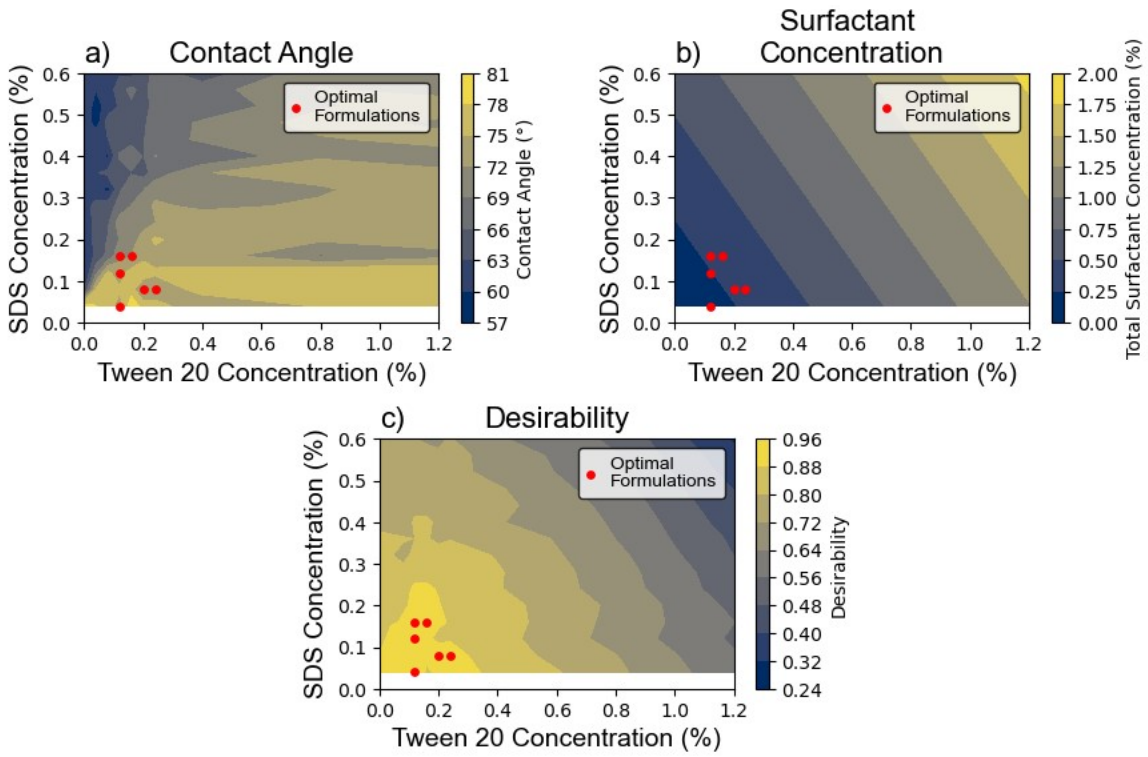

Figure S7. Contour plots of contact angle (a), total surfactant concentration (b), and desirability score (c) across the SDS 94%-Tween 20 formulation space. Red dots indicate Optimal formulations.