DATA FOR *PROGRESS*

PRIORITIES USA

MaxDiff for Message & Policy Testing

January 2022

Background

The Democratic ecosystem relies on a combination of polling, domain expertise, and message testing to develop effective messaging strategy. In order to run winning campaigns, Democrats need an efficient, resilient, and reliable measure of voter preferences across a variety of messages and policies. Over the course of Fall 2021, Priorities USA and Data for Progress worked to verify the efficacy of the MaxDiff (Best-Worst Scaling) question format for political-message testing.

Our research shows that MaxDiff is an effective and simple method of ranking messages, policies, or other discrete elements by capturing the preferences of voters implicitly without asking them to rank the elements themselves. The MaxDiff format works well against traditional measurement tactics like ordinal ranking, Likert scaling, and in-survey Randomized Control Trials (RCTs).

As the 2022 elections approach, movement and electoral partners need clear guidance on which messaging to use for certain issues. MaxDiff offers a powerful solution by answering the question: "If I have to talk about X, what should I say?"

Build Back Better

EXPERIMENTS

In the fall of 2021 Data for Progress conducted an initial MaxDiff poll of likely voters to measure their reactions to messages in support of the Build Back Better agenda. We repeatedly showed respondents 10 random messages, then asked them to rate which messages they preferred most and least each time.

This data is then modeled by DFP using a hierarchical Bayesian model, accounting for basic demographic features of each respondent, to score them with a utility score: that measures the difference between the chance that a message is their most preferred or their least preferred. You can see these scores in the table below.

To verify our MaxDiff rankings, we chose three high-quality measures of message testing and ranking to compare our results: Blue Rose (a reputable progressive analytics firm), a traditional Likert scale (asking respondents to assess how convincing they found each message), and finally the gold-standard for in-survey message testing: a well-powered randomized control trial (RCT) to test the best, middle, and worst messages conducted by Data for Progress.

As the table demonstrates, our RCT results, Blue Rose replication, and Likert scale design tests have validated the ordinal ranking of the MaxDiff utility scores and show that a MaxDiff can capture similar preferences.

In contrast with the traditional methods, the MaxDiff design benefits from a few inherent advantages:

- MaxDiff removes known response biases associated with Likert scales[1,2]
- MaxDiff requires a much smaller sample size than a traditional randomized control trial to test the same number of messages

- MaxDiff can test a variety of formats, from messages to individual politics
- MaxDiff can handle a much larger set of options than traditional methods allow[3]
- MaxDiff provides insight into support for and backlash against messages, compared to Likert scaling.

The table below outlines the relative rankings of messages tested by method. In general, the ranking cross-validate well, with some exceptions. The Likert fails to capture Republican backlash against messages (e.g. extreme weather) which the MaxDiff is better at capturing (neither method captures it as well as RCT).

DFP MaxDiff Rank	Blue Rose Rank	DFP Likert Rank	DFP RCT Rank
Middle Class (+7)	Middle Class	Extreme weather (55%)	
Jobs (+4)	Manufacturing	Healthcare (54%)	Jobs (+5%)
Manufacturing (+4)	Jobs	Jobs (52%)	
Healthcare (0)	Healthcare	Manufacturing (52%)	Healthcare (+2%)
Poverty (-1)	Poverty	Middle class (50%)	
Pandemic (-2)	Pandemic	Poverty (49%)	
21st Century (-4)	21st Century Challenge	Pandemic (48%)	
Extreme Weather (-5)	Global Leader	21st Century (47%)	
Global Leader (-15)	Extreme Weather	Global leader (46%)	
China (-23)	China	China (44%)	China (+0%)

MESSAGES TESTED

1. **Jobs (Best):** We need to pass this plan because it invests in creating thousands of good paying, blue collar jobs that all American workers will benefit from.
2. **Healthcare (Middle):** We need this investment so Americans won't have to worry about burdensome healthcare costs, especially for aging parents.
3. **China (Worst):** The investments will stop America from falling behind foreign competitors like China.
4. **Middle Class:** The investments will grow our middle class, secure a future for our children, and make the American dream affordable again.
5. **Pandemic:** American workers were devastated by the global pandemic. These investments will make sure that our nation is never unprepared again.
6. **Extreme Weather:** As extreme weather conditions become more common, we need these investments to upgrade and safeguard our electric grid.
7. **Poverty:** We shouldn't have to live in fear of poverty. By investing in American families, this plan ensures that we can cut poverty all over our country.

8. **Manufacturing:** This plan invests billions of dollars to rebuild American manufacturing and create a healthy economy.

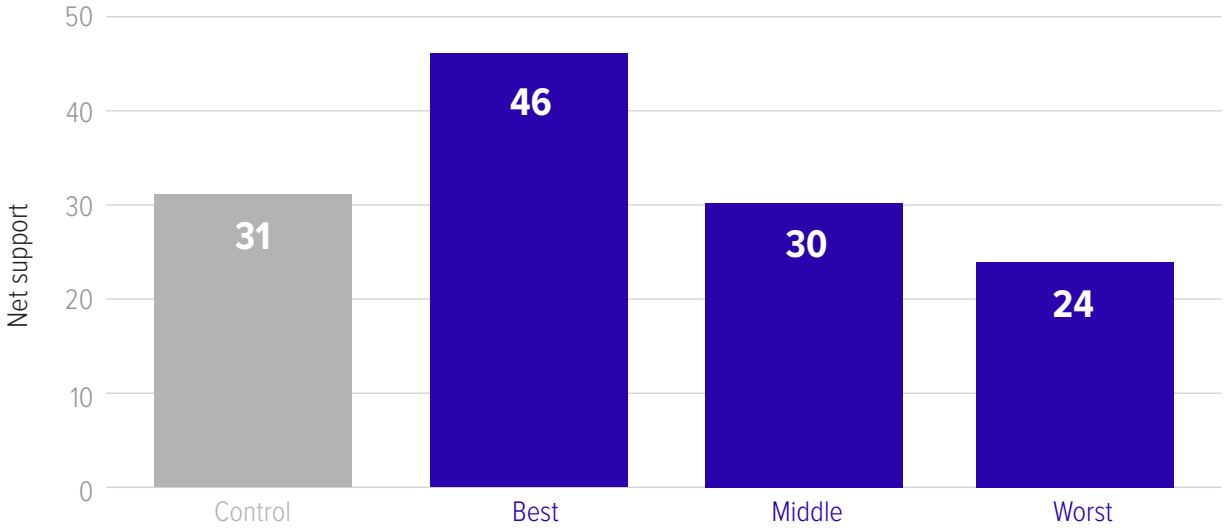
9. **21st-Century Challenge:** Americans overcome every challenge they face. Investing in our people will let us build back better and reach the finish line in the 21st century.

10. **Global Leader:** With these investments, America will continue to be a global leader that no one should ever bet against.

Climate Change Policy

Data for Progress also conducted a second study on MaxDiff to verify our findings using an entirely different policy domain. In this case, we tested eleven core climate policy proposals to see which voters preferred the most and least.

As with Build Back Better messaging, the ordinal ranking of climate policies produced by our MaxDiff survey was consistent with a subsequent RCT with different respondents.



MESSAGES TESTED

1. **Weatherizing (Best):** Weatherizing homes, buildings, and schools to withstand extreme weather and natural disasters

2. **Technology:** Investing in the research and development of new American-made energy technologies

3. **Wind and Solar:** Building new renewable energy projects including wind and solar power

4. **Local Investment:** Providing grants to state and local governments to upgrade local power infrastructure

5. **Climate Corps (Middle):** Creating a Civilian Climate Corps to employ Americans in public works, infrastructure, and conservation projects

6. **Most Affected:** Directing 40% of climate and clean energy investments to low-income communities and communities of color
7. **Pollution:** Penalizing oil and gas producers who create excess amounts of pollution
8. **Tax Loophole:** Eliminating government tax incentives and subsidies for fossil fuel companies
9. **Clean Energy Standard:** Setting a standard to achieve a 100% clean energy electricity grid by 2035
10. **Electric Vehicles:** Making electric vehicles more affordable for consumers
11. **Phase Out Coal (Worst):** Phasing out all coal power plants by 2050

Conclusion

We are confident that MaxDiff offers a quick, powerful, and cost-effective method for ordinally ranking messages and policies by their persuasive effects among voters. In multiple experiments and settings, our format and scoring method validates well against the existing approaches to message testing.

We are excited to continue experimenting with new formats so our partners have access to rapid and robust solutions for identifying the most effective messages to share with voters.

Methodology

All Data for Progress surveys were conducted using web panel respondents. Samples were weighted to be representative of likely voters by age, gender, education, race, and voting history. The surveys were conducted in English.

The margin of error on each survey is between ±2 and ±3 percentage points. Data for Progress interviewed a total of 8,120 likely voters for this study.

DFP uses hierarchical Bayesian modeling to generate scores for voters based on their responses to MaxDiff questions.

Blue Rose uses advanced machine learning and Bayesian techniques to estimate the persuasive and backlash effects of messages among voters.

ENDNOTES

1. Best-worst scaling improves measurement of first impressions
2. Asking About Attitude Change
3. When Choice is Demotivating

COVER PHOTO
Andy Feliciotti/Unsplash