Appendix 2: MTurk Pilot Experiment

Experimental Design

A pilot version of the experiment was run in August 2015 via Amazon's Mechanical Turk (MTurk) service.[1] The design is similar to that of the experiment described in the main text, with four main differences. First, the pilot unfolds all in one stage, giving respondents information about the advisers' statement, the president's decision, and the outcome, and then asking about approval of the president's decision. Second, the pilot omits the discussion of expected armed resistance and holds actual costs constant at zero casualties. Following Kertzer and Brutger (2016), in all conditions in which the president sends troops, I specified that the US did not lose any troops (in addition to holding the outcome of the conflict constant across all conditions, with the attacking country controlling 20 percent of the territory regardless of whether or not the president ultimately sends troops). As Kertzer and Brutger point out, specifying zero casualties in the send-troops conditions holds the costs constant across all conditions, rendering all the conditions logically equivalent, though not lexically equivalent. This design choice is likely to bias support for sending troops upward, however. The experiment reported in the main text adds the expectation of "significant armed resistance" as well as casualties in the send-troops conditions, to show that the adviser cues matter even in the face of expected and actual costs.

A third difference is sample size, which is large (1,690) but still only just over half that of the SSI study reported in the main text. Fourth, consistent with the literature on MTurk samples (e.g., Berinsky, Huber, and Lenz 2012), the MTurk pilot study skews more liberal than the SSI study reported in the main text (difference in means for both party identification and ideology show the MTurk study to have significantly more liberal and Democratic respondents). Although the study is primarily interested in treatment effects—whether advisers can affect approval—the more liberal sample in the pilot study may also affect baseline preferences for staying out of a conflict.

The pilot study was run in two stages. Experiment 1, run on a sample of 953 MTurk respondents, focused only on "costly speech" conditions where the adviser goes against expectations (e.g., a hawkish adviser opposing force, or a dovish adviser advocating force). I then ran a follow-up experiment, Experiment 2, with the "cheap talk" conditions, using a sample of 737 MTurk respondents. Tests indicate no statistically significant differences in baseline approval across the two experiments; I subsequently pooled the two experiments, resulting in a sample of 1,690 respondents with five speech conditions (baseline, two costly speech conditions, and two cheap talk conditions), in addition to variation in the president's party and the decision about whether to use force. Because both experiments had a baseline condition, the pooled results have double the number of respondents in the baseline condition.

[1] Although MTurk samples are not nationally representative, they are an economical and efficient recruitment method that still offers advantages in terms of representativeness as compared with other convenience samples, and comparable results in terms of treatment effects (see Berinsky, Huber, and Lenz 2012).

The basic structure of both (now-pooled) pilot experiments is very similar to the main experiment.[2] After the same introductory prompt described in the main text, respondents read that "A country sent its military to take over a smaller neighboring country. The country that has been attacked is important to U.S. economic and security interests." They are then told that the "The U.S. president, who is a [Democrat | Republican], debated extensively with his advisers about whether to send the military to push back the invaders, or stay out of the conflict. The attacking country continued to invade." Respondents are then randomly assigned to an adviser speech condition—baseline, one of the two cheap talk conditions, or one of the two costly speech conditions, totaling the same five conditions in the main experiment. Then respondents learn the president's decision to either send troops or stay out of the conflict; the vignette notes whether this decision came with the support or opposition of the hawkish or dovish adviser. As discussed above, the final line holds constant the outcome of the conflict. If the president did not send troops, the vignette reports that "The conflict ended with the attacking country taking control of 20 percent of the contested territory." If the president sent troops, the vignette ends by specifying that "The U.S. did not lose any troops in the conflict and the conflict ended with the attacking country taking control of 20 percent of the contested territory." Following the vignette, respondents are asked if they approve or disapprove of the president's handling of the situation, with responses ranging from strongly disapprove to strongly approve on a 7-point scale. The pooled experiments result in a 2x2x5 structure, analogous to Stage 2 of the main experiment.

Results

Results are summarized in Tables A1a-b and Tables A2a-b. Table A1a compares presidential approval for the two decisions (staying out vs. sending troops) across three adviser conditions: baseline, when the adviser supports sending troops, and when the adviser opposes sending troops, the same comparisons shown on the right side of Table 1a in the main text. The pattern is quite similar, although approval levels are somewhat higher. At the stay out decision, approval is a highly significant 12 percentage points lower if the adviser supported the use of force than when the adviser opposed it, and more than two-thirds of respondents approve of the president's decision to stay out if the adviser opposed sending troops. At the send troops decision (again, where in this pilot version, casualties are specified at zero), the approval pattern flips, so that if the adviser supported sending troops, approval is (highly significant) 13 percentage points higher than if the adviser opposed it, and now nearly two-thirds approve of the president sending troops. In Table A1b, the conditions are combined to show that approval is 12 percentage points higher when the president acts in accordance with the adviser's recommendation than when the president acts against the adviser (again, a highly significant difference). The substantive pattern and effect sizes are comparable to those in the main text (Tables 1a and 1b).

Tables A2a and A2b show the results broken down by party and adviser type (analogous to Tables 3a and 3b in the main text). These results should be taken with caution because cell sizes become small in many cases (with approximately 50-70 respondents in some adviser speech conditions). The results for a Democratic president, in Table A2a, suggest that hawkish

[2] I focus here only on the vignette and subsequent approval questions for simplicity; the full survey contained a battery of political knowledge, demographic, and dispositional questions (with the order of the vignette and the dispositional battery randomized), as well as other follow-up questions about the scenario.

advisers have a significant effect on approval for both the stay out and send troops decision; unlike in the main study, dovish advisers have a somewhat significant effect on approval for a Democratic president who sends troops (though again, casualties are held constant at zero in this experiment).

For Republican presidents, Table A2b shows larger effects for obtaining a dovish adviser's support, although the effect is only somewhat significant at the stay out decision and does not reach significance at the send troop decision. The results for Republican presidents, however, are somewhat clouded by the effect of hawkish advisers, whose speech either in support or opposition to sending troops leads to a somewhat significant drop in approval when the president ultimately sends troops (15 percentage points if the hawkish adviser supported troops and 13 percentage points if the hawkish adviser opposed). At the stay out decision, there is similarly little difference between the effect on approval of a hawkish adviser supporting or opposing force, and both conditions are nearly identical to the high (70%) approval for staying out in the baseline condition. Perhaps in this more liberal sample, hawkish advisers are mistrusted, and the decision to send troops or not swamps the effects of hawkish adviser speech. As mentioned, however, dovish advisers do influence approval for Republican presidents and the endorsement of a dovish adviser at least maintains approval at 65%, nearly identical to the baseline level for a Republican who fights.

Overall, the basic contours of the results are similar, especially where presidents wish to play to type. Democratic presidents who wish to play to type—i.e. stay out—benefit from the endorsement of a hawkish adviser, and can get approval as high as 69% if staying out comes with a hawkish adviser's opposition to using force. Republican presidents who wish to play to type—i.e. to send troops—would do well to have their hawkish advisers stay silent, and the endorsement of a dovish adviser can at least prop up support.

Table A1: MTurk Pilot Study,
Effect of Adviser Statements on Presidential Approval

(a) Effect of Adviser Support and Opposition

	Approval if Stay Out		Approval if Send Troops		Effect of Troops	
Baseline: No Speech	63%		65%		+ 2%	
Adviser Supports Troops	56%	*	66%		+ 11%	***
Adviser Opposes Troops	68%		54%	***	-14%	***
Adviser Support vs. Opposition	- 12%	***	+ 13%	***	+ 25%	***

(b) Effect of Action after Adviser Statements

	Approval		Δ from Baseline	
Baseline: No Speech	64%			
Action with Adviser Rec	67%		+ 3%	
Action against Adviser Rec	55%		- 9%	***
Adviser Support vs. Opposition	+ 12%	***		

Note: Percentages are rounded. Asterisks denote the following p-values: *<=.10, **<=.05, and ***<=.01 (two-tailed test).

Table A2a:
Presidential Approval for Democratic Presidents, by Decision and Adviser Statement, MTurk Pilot Study

Democratic President	Stay Out	Δ from Baseline		Send Troops	Δ from Baseline		Swing (Troops-Stay Out)		Δ from Baseline	
No Speech	55%			67%			+ 11%	*		
Hawk Supports Troops	49%	- 6%		68%	+ 1%		+ 18%	**	+ 7%	
Hawk Opposes Troops	69%	+ 13%	*	53%	- 14%	**	- 16%	**	- 27%	***
Hawk Speech Comparison	- 19%	**		+ 15%	*		+ 34%	***		

Democratic President	Stay Out	Δ from Baseline		Send Troops	Δ from Baseline		Swing (Troops-Stay Out)		Δ from Baseline	
No Speech	55%			67%			+ 11%	**		
Dove Supports Troops	54%	- 1%		77%	+ 11%	*	+ 23%	***	+ 12%	
Dove Opposes Troops	65%	+ 10%		59%	- 8%		- 6%		- 17%	
Dove Speech Comparison	- 11%			+ 18%	**		+ 29%	**		
Difference in Hawk swing vs. dove swing (Difference-in-Differences)	- 8%			- 3%			+ 5%			

Note: Percentages are rounded. Asterisks denote the following p-values: *<=.10, **<=.05, and ***<=.01 (two-tailed test).

Table A2b:
Presidential Approval for Republican Presidents, by Decision and Adviser Statement, MTurk Pilot Study

Republican President	Stay Out	Δ from Baseline	Send Troops	Δ from Baseline	Swing (Troops-Stay Out)	Δ from Baseline
No Speech	70%		64%		- 6%	
Hawk Supports Troops	71%	+ 1%	49%	- 15% *	- 21% **	- 16%
Hawk Opposes Troops	68%	- 1%	51%	- 13% *	- 18% **	- 12%
Hawk Speech Comparison	+ 2%		- 1%		- 4%	

Republican President	Stay Out	Δ from Baseline	Send Troops	Δ from Baseline	Swing (Troops-Stay Out)	Δ from Baseline
No Speech	70%		64%		- 6%	
Dove Supports Troops	54%	- 16% **	65%	+ 1%	+ 11%	+ 17% *
Dove Opposes Troops	70%	+ 0%	54%	- 10%	- 16% *	- 10%
Dove Speech Comparison	- 16% **		+ 11%		+ 27% **	
Difference in Hawk swing vs. dove swing (Difference-in-Differences)	+ 18%		- 12%		- 30% *	

Note: Percentages are rounded. Asterisks denote the following p-values: *<=.10, **<=.05, and ***<=.01 (two-tailed test).