

# INVESTIGATION OF VARIABILITY IN SEVERITY RESPONSE SET COMPOSITION ACROSS PROs: DATA FOR INCREASING CONSISTENCY, FACE VALIDITY, AND COMPREHENSIBILITY IN INSTRUMENT DEVELOPMENT



Rebecca Israel, Tim Poepsel, Allyson Nolde, Chryso Hadjidemetriou, Oliver Delgaram-Nejad, Payton Ramsey, Rachael Browning, Elizabeth McCullough, Shawn McKown

## INTRODUCTION:

Ordinal scales with discrete response options commonly appear in PROs to assess various dimensions (e.g., frequency, severity) of health conditions or treatments.

Previous research<sup>1,2,3</sup> has demonstrated negative effects on patient comprehension and translatability related to increased response set size (e.g., 7 vs 5 response options) and resulting poor conceptual spacing between options.

Other potential problems for patients are:

1. variability within and across instruments in the composition of response sets assessing a particular dimension such as frequency or severity
2. and poor response set balance - where "balance" is defined as equivalent numbers of response options describing high and low target dimension values. Ideally, ordinal response sets should have a clear internal ordering, and equal space on either side of a scale midpoint, so that patient responses aren't biased towards a particular scale end

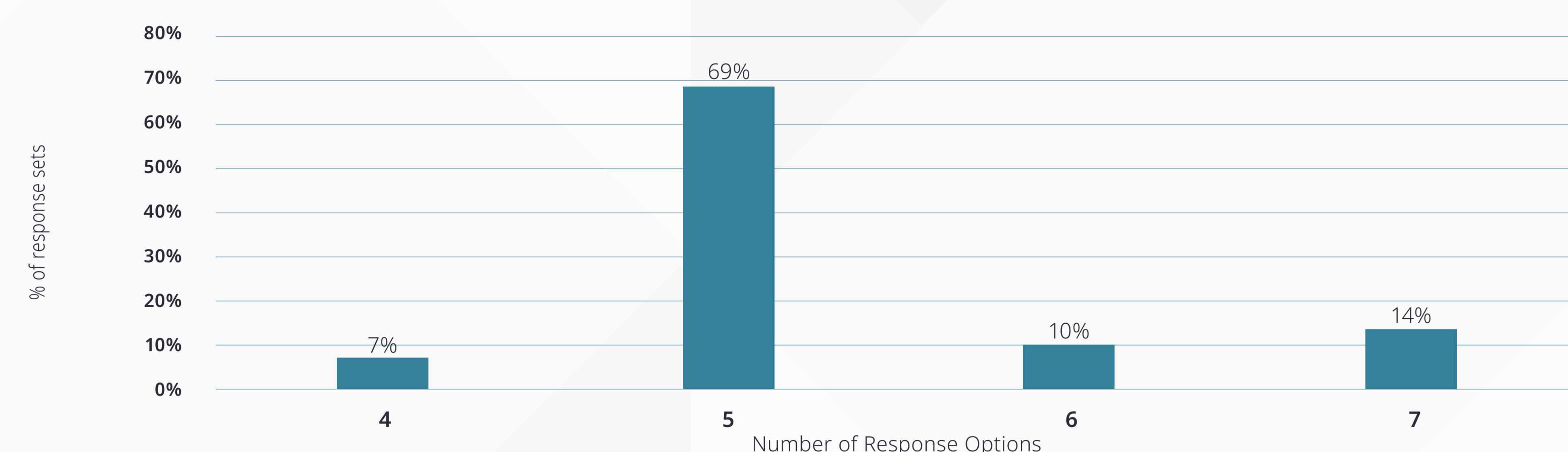
We aimed to characterize these potential problems by examining the range of variability in response set composition and balance in existing PROs, specifically for the **severity** dimension.

## METHODS:

We reviewed ~2000 PROs from previous linguistic validation projects

- finding **42 unique ordinal severity response sets**
- **57 unique severity response options**
- each containing between **4 and 7 response options** (See Figure 1 for distribution)
- 83% of sets had an **odd number** of response options and thus a **natural midpoint**

Figure 1: Response Set Size Distribution

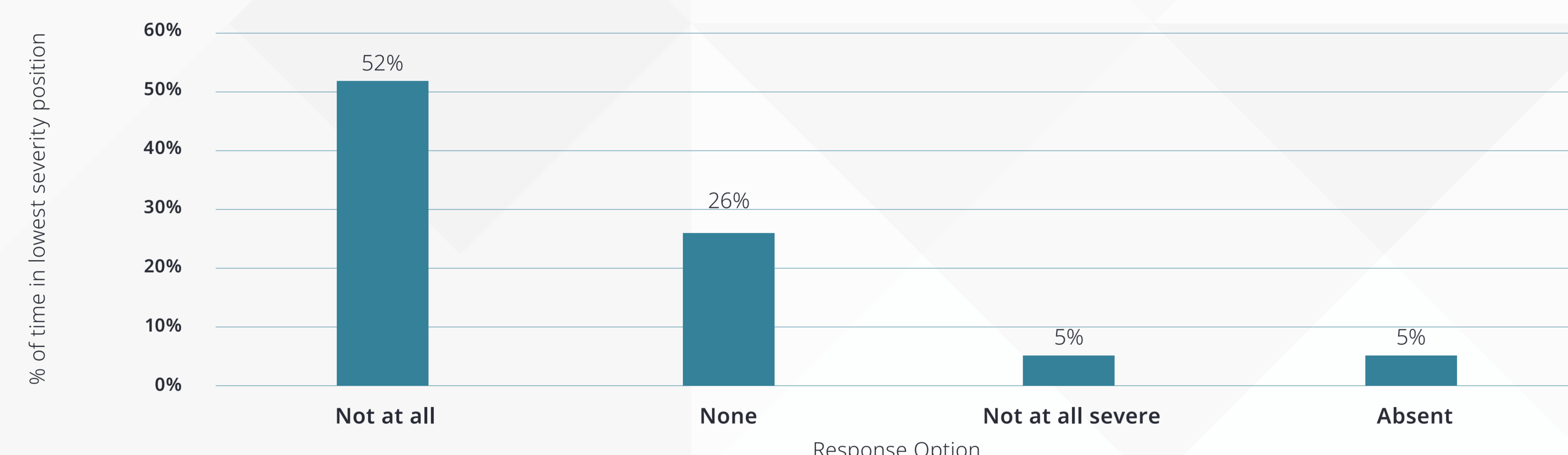


## RESULTS:

Across all 42 sets, 9 response options appeared in the **lowest severity position** (See Figure 2 for the top four response options by frequency of appearance in the lowest severity position). Other "lowest severity" response options are shown in Table 2.

- "Not at all" was the most common "lowest severity" option, appearing there 52% of the time

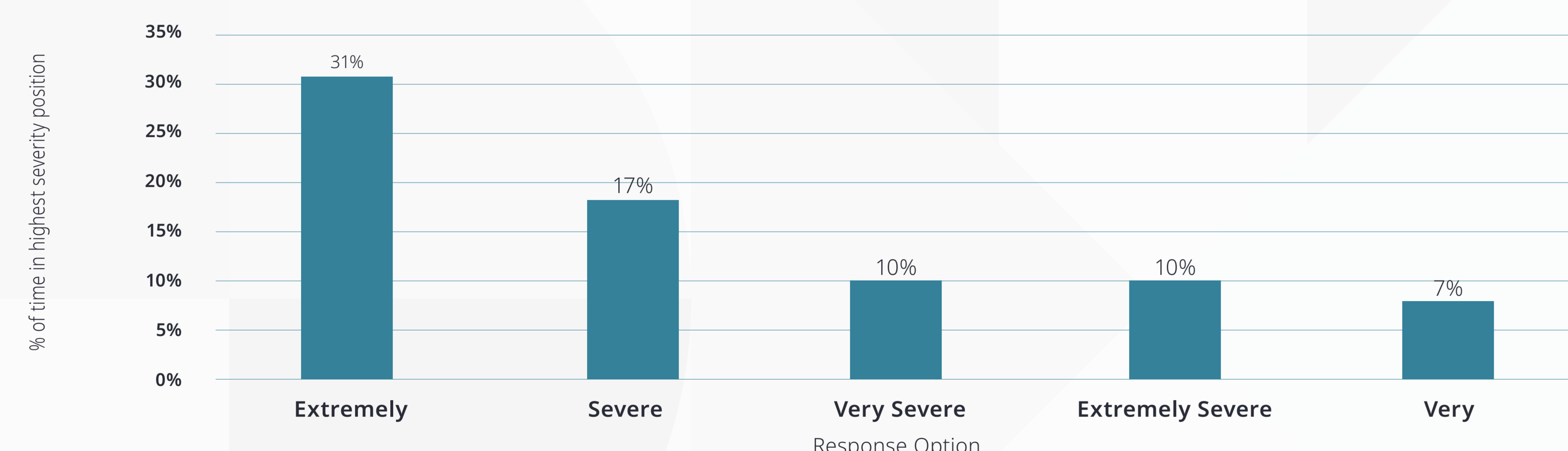
Figure 2: Most Common 'Lowest Severity' Response Options



15 response options appeared in the **highest severity position** (See Figure 3 for top 5 response options by frequency of appearance in the highest severity position). Other "highest severity" response options are shown in Table 2.

- "Extremely" was the most common "highest severity" option, appearing there 31% of the time

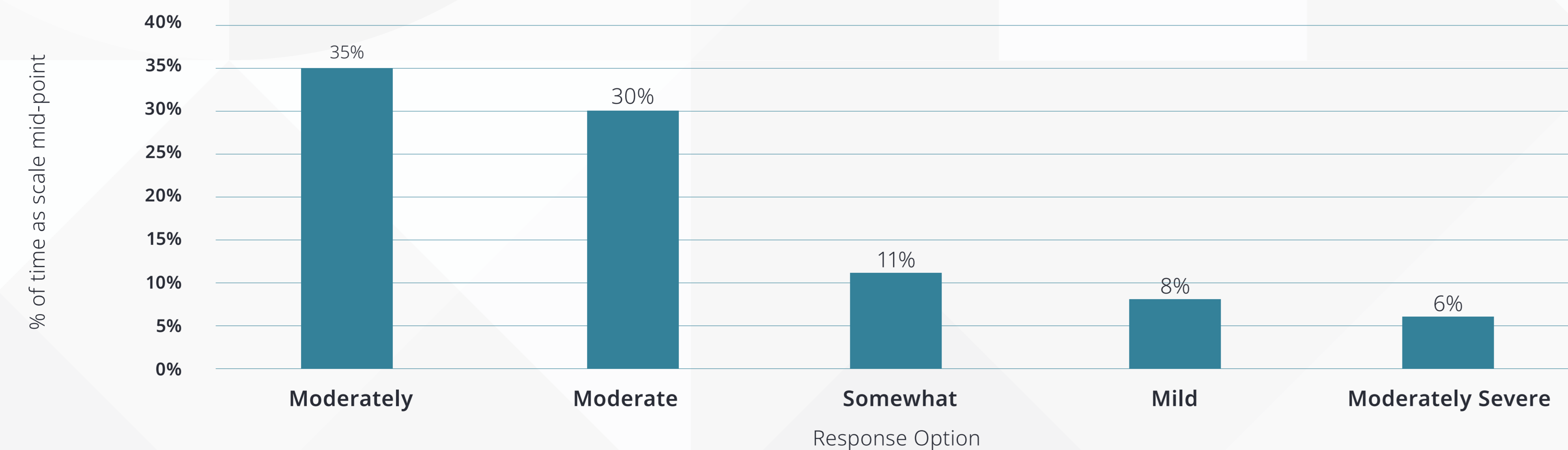
Figure 3: Most Common 'Highest Severity' Response Options



9 response options appeared as **scale mid-points** (for odd-numbered response sets; See Figure 4 for top 5 response options by frequency of appearance as scale midpoints). Other scale midpoints are shown in Table 2.

- "Moderately" (35%) and "Moderate" (30%) were the most common severity scale mid-points, appearing there a combined 65% of the time

Figure 4: Most Common Severity Midpoint Response Options



35 response options appeared outside of the highest, lowest and scale midpoint positions, to represent both high and low severity response options (See Table 2 for a listing of these 'other' response options).

Qualitative analysis by instrument design experts found 33% of severity sets were unbalanced. See Table 1 for examples of unbalanced response sets, and Figures 5 and 6 for a more detailed analysis of their distribution. 5-point response sets, which were the most common in our sample (N=29), were unbalanced 17% of the time. 4-point response sets were unbalanced 100% of the time; 6-point response sets were unbalanced 75% of the time, and 7-point response sets were unbalanced 50% of the time.

Figure 5: Percentage of Balanced Response Sets

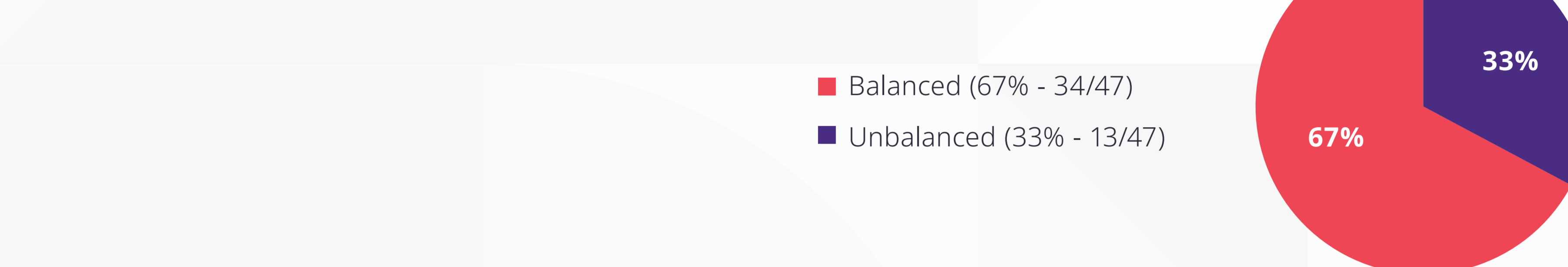


Figure 6: Unbalanced Response Sets by Number of Response Options in Set

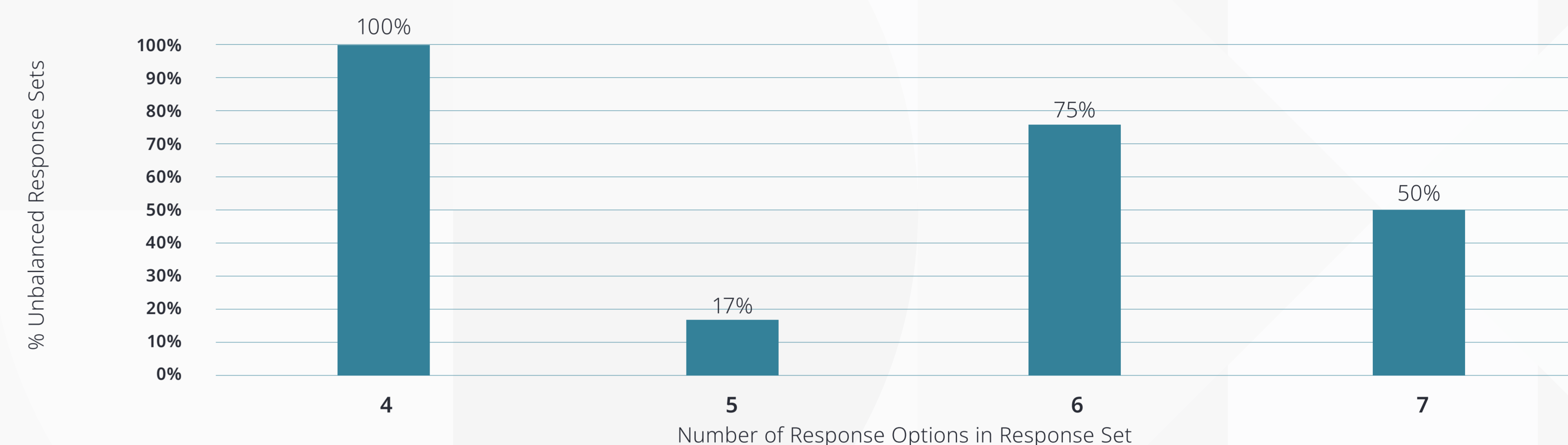


Table 1: Examples of Unbalanced Response Sets

Example Set 1	Example Set 2	Example Set 3
Not at all	Absent	None
A little bit	Minimal	Very mild
Somewhat	Mild	Mild
A good bit	Moderate	Moderate
Quite a bit	Severe	Severe
Very much		
Extremely		

Table 2 shows all response options in this data set by the frequency of their appearance in a certain scale position (e.g., highest severity, lowest severity, etc.), and critically shows where positional overlap occurs for specific response options. Response options with positional overlap are denoted by highlighting in the table. We found positional overlap for 19% (11/57) of the response options studied here, and furthermore, those overlapping response options account for 43% (94/221) of the response option appearances in our data set, suggesting high potential for confusion regarding the severity level that certain response options signal. Of the 11 response options with positional overlap, 6 (55%) of them are among the most frequent for their respective scale position ("very severe", "severe", "moderately", "moderate", "mild", "somewhat" - see Figures 3 and 4), again suggesting the potential for patient confusion and ambiguity of signaling for common and critically positioned response options.

Table 2: All Response Options by Frequency of Appearance in Response Set Positions

Response Option	Highest Severity	Mid-point	Lowest Severity	Other
Very severe	4			2
Extremely severe	4			
Severe	7			10
Extremely	13			
Completely	1			
Overwhelmingly	1			
Very	3			6
A great deal	2			1
A very great deal	1			
Extreme	1			
Worst possible	1			
A lot	1			1
Intolerable	1			
Very much	1			
A whole lot	1			
None			11	
Not at all severe			2	
No			1	
Not at all			22	
Not			1	
Absent			2	
Not present			1	
Minimal			1	2
A lot better			1	
Moderate		11		5
Moderately severe		2		3
Moderately		13		1
Some		1		
Mild		3		8
Somewhat			4	
Average		1		
No change		1		
A good bit		1		
Mildly severe				1
Slight				2
Marked				1
A little bit				6
Quite a bit				9
Mildly				2
Severely				3
Borderline				1
Markedly				1
A little				8
Some				1
A little severe				1
Somewhat severe				1
Minor				2
Very mild				3
Mostly				2
Quite				2
Fairly				1
Minimally				1
Slightly				5
A little better				1
A little worse				1
Very much				1
Fairly good				1

## CONCLUSIONS

These results show great variability in severity response set composition and the identity of the highest and lowest severity response options, as well as scale midpoints.

These results also show an alarmingly high percentage of unbalanced sets, especially for 4, 6, and 7-point response sets. 5-point response sets are comparatively less unbalanced, but the rate is still fairly high at 17%.

We also show a moderate to high degree of overlap in how individual response options are used, which may contribute to patient confusion regarding their meaning. Positional overlap was a potential confounder for over 40% of occurrences of severity response options in our sample of 42 unique response sets.

These factors may reduce response set comprehensibility, impact PRO function within and across trials, and reduce data quality. Our results demonstrate a need for increased homogeneity of response sets assessing particular dimensions, and more careful selection of response options within a response set, to improve balance and maximize the ability of response options to signal a particular scale position.