Supplemental Material

CBE—Life Sciences Education

Deane *et al.*

Throughout the process, the opening element of the question remained the same: **Carefully read four potential outcomes of your experiment (listed below, 1–4). Outcome 1: None of the other 4 species have any effect on cheatgrass growth rate. Outcome 2: Only 1 other species increases cheatgrass growth rate. Outcome 3: Only 1 other species decreases cheatgrass growth rate. Outcome 4: All the other 4 species decrease cheatgrass growth rate.**

First Stage (original question and answers)	**Example Student Responses (used to shape development)**
Which of the following outcomes would be most scientifically useful? *a) Outcomes 1 or 4* *b) Outcomes 3 or 4* *c) Outcomes 2, 3 or 4* **d) Outcomes 1, 2, 3 or 4**	"…I'd choose c) because we're talking about crop management here so it'd be more useful to find a species that would stop the weed." [***Note: There were many variations of this response***] "I'd choose b) because I want to find a way to stop the grass spreading so I want to know which ones would do that." [***Note: There were a few variations of this response***] "I think it would be useful to state in the answers that if I choose one (answer), that means these would be more useful than the others as right now I don't know if I'm comparing them against each other or not."

To address the issues raised from student interviews at the first stage of this question, (for the second stage) we tried to work on the terminology in the question stem to make it clear that we were asking about the value of conducting research (and that any result would be useful because it would provide new information), and not about which result would be best in terms of preventing cheatgrass from spreading. We also incorporated relative comparisons into the answer stems.

Second Stage (question and answers)	**Example Student Responses (used to shape further development)**
Which of the following outcomes would be most useful in terms of advancing knowledge? *a) Outcomes 1 or 4 (more useful than outcomes 2 or 3)* *b) Outcomes 3 or 4 (more useful than outcomes 1 or 2)* *c) Outcomes 2, 3 or 4 (more useful than outcome 1)* **d) Outcomes 1, 2, 3 or 4 (all equal)**	"…Go with c) because knowledge about stopping cheatgrass is what you want." [***Note: There were still a few variations of this response***] "I think it could be a) actually. If we know flat out that either none or all of the species stop the cheatgrass then that would advance my knowledge more than the other options. Like… even if none of them work that's more useful for me to stop the cheatgrass as I can try other species instead. [***Note: When pushed to explain this more: "Yeah, so if I'm trying to 'advance knowledge' I'm just looking for the science solution to the problem."***]

To address the issues raised from student interviews at the second stage of this question, (for the third stage), we worked more on the terminology in the question stem to make it clear that we were asking about the purpose of conducting the experiment (and that any result would be useful because it would provide new information), and not about which result would be best in terms of preventing cheatgrass from spreading.

Third Stage (question and answers)	Example Student and Expert Responses (used to confirm this final version as suitable)
Which of these outcomes would **most** advance scientific understanding about the interactions between these species?	"I know more about the interactions between the species if they either all stop the weed or if none of them do because then they all have the same impact, so [answer] a)"
a) Outcomes 1 or 4 (would advance understanding more than outcomes 2 or 3)	"I choose b). This improves my understanding as I then understand something positive about the interactions. A negative interaction – like if the cheatgrass improves with the [other species] doesn't improve my understanding."
b) Outcomes 3 or 4 (would advance understanding more than outcomes 1 or 2)	"[Answer] c) because it cant be Outcome 1. If none of the other species affect it I wouldn't have advanced the understanding because I still wouldn't know which species affect it."
c) Outcomes 2, 3 or 4 (would advance understanding more than outcome 1)	**"It's [answer] d) because it doesn't matter what the other species do to the cheatgrass. As long as I've done the experiment I will then know whether they affect it and that would advance it as it would be new information." [Note: There were many frequent variations of this expert-like thinking]**
d) Outcomes 1, 2, 3 or 4 (all would advance understanding equally)	*** For all of the above examples, the responses validate the question, because students and experts were choosing the answers that matched their reasoning (whether this was considered expert-like, or non–expert-like).**