Table of Contents

# Assessment Bank Export

Generated: 2025-11-01T03:27:43.178Z

Schemas: 5 • Items: 18

## Table of Contents

* [AlternativeExplanationGeneration](#schema-alternativeexplanationgeneration)
* [BiasDirectionOpen](#schema-biasdirectionopen)
* [BiasDirectionSequential](#schema-biasdirectionsequential)
* [FermiPopulationCity](#schema-fermipopulationcity)
* [SelectionEffectIdentification](#schema-selectioneffectidentification)

## Schema: AlternativeExplanationGeneration

* Description: Assesses the ability to generate distinct, plausible alternative explanations for an association.
* Ability: causal.aeg

### AEG\_Marriage\_Happiness\_T1

Stem: Married people report higher levels of happiness, on average, than unmarried people. Give one reason why this might be true, other than marriage causing happiness.

* Scenario:
  + A: being married
  + B: higher levels of happiness
* Theme Registry:
  + reverse\_causation\_selection — happier people get married; unhappy people divorce; optimism
  + ses\_financial\_stability — financial stability; job security; wealth
  + personality\_traits — agreeableness; conscientiousness; emotional stability
  + health\_status — better physical health; mental health; healthy lifestyle
  + social\_support\_networks — family support; community involvement; social integration
* Too General: personality, background, lifestyle, culture, environment
* Examples:
* RunsThroughA:
  + Marriage provides companionship and emotional support, which makes people happier. ( This is A->B. Companionship/support is the mechanism by which marriage causes happiness. )
  + Married couples share expenses, so they have less financial stress, leading to more happiness. ( This is A->Mediator->B. Shared expenses/economies of scale (Mediator) are a direct consequence of being married (A). )
  + Being married reduces loneliness, and loneliness causes unhappiness. ( This is A->Displacement->B. Marriage (A) displaces loneliness, leading to happiness (B). )
  + Married people have a partner to help them through tough times, which increases happiness. ( This is A->B. Having a partner is the mechanism by which marriage causes happiness. )
* Good:
  + People who are already happier and more emotionally stable are more likely to attract a partner and get married. ( This correctly identifies reverse causation (selection). )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between marriage and happiness?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the explanation you're giving?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between being married and higher levels of happiness.
  + Please focus on explaining the association described in the question.

### AEG\_MBA\_Salary\_T1

Stem: People who have an MBA degree earn significantly higher salaries, on average, than those who do not. Give one reason why this might be true, other than the degree (or the education) causing the higher salary.

* Mutually Exclusive:
  + Group: income\_determinants
  + With: AEG\_SelfConfidence\_Income\_T1
* Scenario:
  + A: having an MBA degree
  + B: higher salaries
* Theme Registry:
  + prior\_ability\_experience — more work experience; academic achievement; cognitive ability
  + ambition\_drive — conscientiousness; willingness to negotiate salary; strong work ethic
  + ses\_network\_capital — family wealth; pre-existing professional networks; privileged background
  + industry\_sorting — work in finance; tech industry; choice of career field
  + company\_sponsorship — employer paid for MBA; sponsored by high-paying firm; identified as high-potential employee
* Too General: background, personality, where they work, selection effects
* Examples:
* RunsThroughA:
  + The MBA teaches valuable business skills and management techniques that employers pay more for. ( This is A->B (Human Capital Theory). The skills learned during the MBA (A) cause the higher salary (B). )
  + The MBA degree acts as a signal of competence to employers, so they are willing to pay them more. ( This is A->B (Signaling Theory). The degree itself causes the higher pay because of how employers interpret the credential. )
  + You gain access to the business school's alumni network during the program, which helps you get better jobs. ( This is A->Mediator->B. The network gained during the MBA (Mediator) is a direct consequence of pursuing the degree (A). This must be distinguished from pre-existing networks. )
  + The MBA credential itself qualifies them for high-paying management roles that others can't get. ( This is A->B (Credentialism). The degree causes the higher salary by acting as a prerequisite. )
  + The MBA gives people the confidence to ask for higher salaries. ( This is A->Mediator->B. The confidence (Mediator) is caused by the degree (A). )
* Good:
  + People who are highly ambitious are more likely to pursue an MBA and also more likely to work harder and negotiate higher salaries. ( This correctly identifies a common cause (ambition). )
  + Individuals from wealthy backgrounds can afford the tuition for an MBA and also have family connections that lead to better jobs. ( This correctly identifies a common cause related to SES and networks. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between having an MBA and higher salaries?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between having an MBA degree and higher salaries.
  + Please focus on explaining the association described in the question.

### AEG\_PetStress\_T1

Stem: Adults who own a pet report lower levels of stress, on average, than those who do not. Give one reason why this might be true, other than pet ownership reducing stress.

* Mutually Exclusive:
  + Group: stress\_relief\_lifestyle
  + With: AEG\_TeaDrinking\_Stress\_T1, AEG\_SocialMedia\_Anxiety\_T1
* Scenario:
  + A: owning a pet
  + B: lower levels of stress
* Theme Registry:
  + reverse\_causation\_capacity — life stability; better mental health; more free time
  + ses\_resources — financial stability; better housing; predictable schedules
  + personality\_sociality — extroversion; agreeableness; strong social support
  + age\_lifestage — older adults; retirees
* Too General: lifestyle, environment, personality, background, demographics
* Examples:
* RunsThroughA:
  + Walking a dog provides exercise, and exercise reduces stress. ( This is A->Mediator->B. Dog walking (Mediator) is a direct consequence of owning the pet (A). This relies on the prohibited pathway. )
  + Pets provide companionship that lowers stress. ( This is A->B. Companionship is the mechanism by which pets reduce stress. )
  + Pets make you happy. ( This is A->B. Happiness is a mediator for reduced stress. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between pets and lower stress?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between owning a pet and lower stress levels.
  + Please focus on explaining the association described in the question.

### AEG\_SelfConfidence\_Income\_T1

Stem: People with higher self-confidence tend to earn higher incomes. Give one reason why this might be true, other than higher self-confidence causing higher income.

* Mutually Exclusive:
  + Group: income\_determinants
  + With: AEG\_MBA\_Salary\_T1
* Scenario:
  + A: higher self-confidence
  + B: higher income
* Theme Registry:
  + reverse\_causation\_success — income boosts confidence; career success builds confidence; wealth increases self-esteem
  + prior\_skill\_competence — intelligence; expertise; track record of success
  + family\_ses\_capital — family wealth; elite education access; social capital
  + personality\_traits — risk-taking personality; assertiveness; emotional stability
  + physical\_attributes — attractiveness; height; beauty bias
* Too General: background, personality, lifestyle
* Examples:
* RunsThroughA:
  + Confident people are more willing to negotiate for higher salaries. ( This is A->Mediator->B. Negotiation (Mediator) is driven by confidence (A) and leads to higher income (B). )
  + People with high self-confidence seem more competent to employers, so they get promoted more often. ( This is A->B. Confidence causes higher income because of how it is perceived by others (signaling/halo effect). )
  + Confident people take more career risks, which often lead to higher rewards. ( This is A->Mediator->B. Risk-taking (Mediator) is driven by confidence (A). )
  + High self-confidence improves performance in job interviews, leading to better offers. ( This is A->B. Confidence improves interview performance which leads to higher pay. )
* Good:
  + Being highly skilled at your job leads to success (higher income) and also makes you feel confident about your abilities. ( This correctly identifies a common cause (skill/competence). )
  + Earning a lot of money might make people feel more confident about themselves over time. ( This correctly identifies reverse causation (Income -> Confidence). )
  + People from wealthy backgrounds might have access to networks that lead to high-paying jobs, and their upbringing might also foster higher confidence. ( This correctly identifies a common cause (SES/Social Capital). )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between self-confidence and higher income?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between higher self-confidence and higher incomes.
  + Please focus on explaining the association described in the question.

### AEG\_SocialMedia\_Anxiety\_T1

Stem: People who use a lot of social media report higher levels of anxiety. Give one reason why this might be true, other than social media causing more anxiety.

* Mutually Exclusive:
  + With: AEG\_PetStress\_T1, AEG\_TeaDrinking\_Stress\_T1
* Scenario:
  + A: using social media a lot
  + B: higher anxiety
* Theme Registry:
  + age\_structure — younger age; college-age
  + isolation\_loneliness — few offline friends; weak local ties; lonely
  + reverse\_coping — avoidant coping; stress leads to more use; anxious people scroll
* Too General: resources, environment, lifestyle, culture, support structures
* Probes:
* Good:
  + Thank you. Can you think of any other reason for the link?
  + Great. Now please try to give another, different explanation.
  + Thanks. Next, please think of another distinct explanation.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about your explanation?
  + Can you be more specific about how that would work?
  + That explanation is a bit broad. Can you give a concrete example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + Could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation.
* NotRelevant:
  + Please focus on how to explain the link described in the question.
* RunsThroughA:
  + Please avoid relying on social media use directly causing anxiety. Try a different kind of explanation.

### AEG\_TeaDrinking\_Stress\_T1

Stem: People who regularly drink tea report lower levels of stress, on average, than those who do not. Give one reason why this might be true, if drinking tea does not actually reduce stress.

* Mutually Exclusive:
  + Group: stress\_relief\_lifestyle
  + With: AEG\_PetStress\_T1, AEG\_SocialMedia\_Anxiety\_T1
* Scenario:
  + A: regularly drinking tea
  + B: lower levels of stress
* Theme Registry:
  + reverse\_causation\_pace\_of\_life — less stressed people drink tea; more free time
  + health\_behaviors\_bias — healthy diet; exercise regularly
  + ses\_resources — financial stability; job security; flexible work hours
  + culture\_geography — cultural background; where they live; social norms
  + age\_lifestage — older adults; retirees; life stage
  + personality\_disposition — conscientiousness; emotional stability
* Too General: lifestyle, environment, personality
* Examples:
* RunsThroughA:
  + Tea contains chemicals like L-theanine which are known to be calming. ( This is A->B. This describes the biological mechanism by which tea reduces stress. )
  + The warmth of the tea is physically relaxing. ( This is A->B. This describes a physical mechanism by which drinking tea reduces stress. )
  + The ritual of brewing and drinking tea forces you to take a break, which reduces stress. ( This is A->Mediator->B. The ritual/break (Mediator) is a direct consequence of drinking tea (A). This relies on the prohibited pathway. )
  + People drink tea instead of coffee, and less caffeine means less stress. ( This is A->Displacement->B. Drinking tea (A) displaces coffee (a stressor), leading to lower stress (B). This relies on A causing the outcome. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between tea drinking and lower stress?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between drinking tea and lower stress levels.
  + Please focus on explaining the association described in the question.

### AEG\_VideoGames\_Aggression\_T1

Stem: Children who play violent video games are more likely to show aggressive behavior in school. Give one reason why this might be true, other than playing violent video games causing aggressive behavior.

* Scenario:
  + A: playing violent video games
  + B: aggressive behavior in school
* Theme Registry:
  + underlying\_traits — impulsivity; sensation-seeking; low empathy
  + home\_environment — exposure to real violence; lack of supervision; family conflict
  + social\_peer\_group — peer pressure; delinquent peers; social norms
  + reverse\_causation\_rejection — aggression leads to social rejection; retreating into games; bullying (as victim)
  + demographics\_gender — being male
* Too General: personality, environment, background, demographics
* Examples:
* RunsThroughA:
  + The games desensitize children to violence, making them more likely to act aggressively. ( This is A->B (Desensitization). This is the mechanism by which games cause aggression. )
  + Children imitate the violent behaviors they practice in the video games. ( This is A->B (Social Learning/Imitation). This is the mechanism by which games cause aggression. )
  + Playing the games makes them frustrated, and this frustration leads to aggression at school. ( This is A->Mediator->B. Frustration (Mediator) is caused by playing the game (A). )
  + Playing video games takes time away from learning social skills, which leads to aggression. ( This is A->Displacement->B. Gaming (A) displaces social learning, leading to aggression (B). This relies on A causing the outcome. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. Can you think of another reason for the link between violent video games and aggressive behavior?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between playing violent video games and aggressive behavior.
  + Please focus on explaining the association described in the question.

### AEG\_WFH\_JobSatisfaction\_T1

Stem: People who work from home report higher job satisfaction, on average, than those who work in an office. Give one reason why this might be true, other than working from home causing higher job satisfaction.

* Mutually Exclusive:
  + Group: wfh\_job\_satisfaction
  + With: BDO\_RemoteWork\_NullEffect\_T1
* Scenario:
  + A: working from home (WFH)
  + B: higher job satisfaction
* Theme Registry:
  + industry\_job\_type — tech industry; creative jobs; high-autonomy fields
  + seniority\_status — management positions; tenure; more control over work
  + personality\_preference — introversion; self-motivated; disciplined
  + reverse\_causation\_trust — trusted by management; high performers allowed WFH; already satisfied employees WFH
  + company\_culture — progressive companies; better management; high-trust organizations
* Too General: lifestyle, background, demographics
* Examples:
* RunsThroughA:
  + Working from home eliminates the stress of commuting, which increases job satisfaction. ( This is A->Mediator->B. Reduced commute stress (Mediator) is a direct result of WFH (A). )
  + WFH allows for better work-life balance, making employees happier with their jobs. ( This is A->B. Work-life balance is the mechanism by which WFH increases satisfaction. )
  + People are more focused and productive at home because there are fewer distractions, which leads to higher satisfaction. ( This is A->Mediator->B. Increased focus/fewer distractions (Mediator) is caused by WFH (A). )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please share only one explanation for now; then we can ask for another.
* Good:
  + Thank you. What's another possible reason for the link between working from home and higher job satisfaction?
  + Great. Now please try to give a different explanation for the association.
  + Thanks. Next, please think of another distinct reason.
* NotDistinct:
  + That is a bit too similar to an explanation you already gave. Please provide a different one.
  + That seems to rely on the same kind of cause as a previous answer. Can you give a different kind of explanation?
  + Ok, this time try to think about explanations of a different type.
* NotSpecific:
  + Could you give a bit more detail about how that would work?
  + Can you be more specific about the mechanism you are describing?
  + That explanation is a bit broad. Can you give a concrete factor or example?
* NotClear:
  + Can you clarify what you mean?
  + I don't quite understand. Can you rephrase that?
  + I didn't follow—could you explain how that would work?
* NotPlausible:
  + Ok, this time try a different kind of explanation for the link.
  + Let's try another angle—please offer a different reason.
* NotRelevant:
  + Please focus on explaining the link between working from home and higher job satisfaction.
  + Please focus on explaining the association described in the question.

## Schema: BiasDirectionOpen

* Description: Assesses recognition that observational null results are ambiguous due to selection biases that can mask benefit or harm; constrained categorical AJ output for dynamic probing and path-dependent scoring.
* Ability: causal.interpretation.null\_effects

### BDO\_CharterSchools\_NullEffect\_T1

Stem: A comprehensive study found that students attending elective public charter schools had no difference in learning outcomes from those attending traditional public schools in the same districts.

What do we learn from this about the actual effect that charter schools in the district had on learning outcomes? Explain your reasoning.

* Scenario:
  + Intervention: Attending a charter school (vs. traditional public school).
  + Outcome: Academic performance.
  + Observed Result: No difference in average performance between the two groups (Null effect).
* Examples:
* MaskedBenefit\_Example:
  + The charter schools might actually be more effective (Benefit). It's possible that students who were struggling in traditional schools were more likely to switch to a charter school seeking help (Negative Bias). If the charter school successfully educated them, it would bring their performance up to the average, resulting in the 'no difference' finding. ( Identifies Negative Bias (struggling students select in) and Masking Logic (effectiveness brings them up to average). )
* MaskedHarm\_Example:
  + The charter schools might actually be less effective (Harm). It's plausible that the families most motivated and resourced are the ones who navigate the application process to get into charter schools (Positive Bias). If the charter schools provided a worse education, it could bring the performance of these high-potential students down to the average. ( Identifies Positive Bias (motivated/resourced families select in) and Masking Logic (ineffectiveness brings them down to average). )
* Neither\_Explained\_Sufficiently\_Example:
  + You can't conclude anything because the families chose the schools themselves; it wasn't randomized. There are too many variables. ( This identifies the lack of randomization but does not provide a specific selection mechanism or explain the masking logic. )
* Probes:
* Both\_Explained:
  + Thank you. That is a comprehensive interpretation of the charter school study.
  + Thanks, that covers the key possibilities regarding the charter schools' effectiveness.
  + Got it. That's a thorough analysis of the comparison.
* MaskedBenefit\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the charter schools were actually *less effective* than the public schools?
  + Thanks. What about the other side? Explain how the results might look this way even if the charter schools were actually worse.
  + Understood. Now, please explain a scenario where charter schools decrease performance, but the study still shows no difference.
* MaskedHarm\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the charter schools were actually *more effective* than the public schools?
  + Thanks. What about the other side? Explain how the results might look this way even if the charter schools were actually better.
  + Understood. Now, please explain a scenario where charter schools improve performance, but the study still shows no difference.
* Neither\_Explained\_Sufficiently:
  + Is the observed 'no difference' sufficient evidence to conclude charter schools are exactly as effective as public schools? Why or why not?
  + Can you think of a scenario where the charter schools *are* different in effectiveness, but the study still shows no difference?
  + Why might the observed results be misleading about the schools' actual effectiveness?
* NotSpecific:
  + Can you be more specific about how that would lead to the observed 'no difference' in academic performance? Please explain the mechanism.
  + That's a bit broad. Can you describe the specific factor and how it might mask the schools' actual effectiveness?
  + Could you explain the full logic? How exactly does the factor you mentioned lead to the observed 'no difference' result?
* NotDistinct:
  + That relies on the same underlying logic as your previous explanation. Please provide a distinctly different possibility for the schools' actual effectiveness.
  + That's too similar to the possibility you already mentioned. Try explaining how the opposite effect could be true.
  + This seems to overlap with an earlier point. Can you describe a different scenario for the schools' impact?
* NotPlausible:
  + Let's try another angle. Please offer a different interpretation of the results.
  + Okay, please provide a different possibility regarding the schools' effectiveness.
  + Try to think of a different reason why the study showed no difference.
* NotClear:
  + I didn't follow your reasoning. Could you rephrase your explanation?
  + Can you clarify what you mean? I don't quite understand.
  + I don't quite follow. Could you explain how that would work?
* NotRelevant:
  + Please focus on interpreting the results of the study regarding charter school effectiveness.
  + Your response seems off-topic. Please focus on the effectiveness of the charter schools based on the study.
  + Please address the prompt: what is your interpretation of the study results?

### BDO\_IntensiveTherapy\_NullEffect\_T1

Stem: A hospital introduced an optional, intensive therapy program for patients recovering from a specific type of injury. But patients in the program had the same recovery time as those who did not.

How should this finding be interpreted when considering the program's effect on recovery time? Explain your reasoning.

* Scenario:
  + Intervention: Optional intensive therapy program (vs. standard protocol).
  + Outcome: Average recovery time.
  + Observed Result: No difference in recovery time between the two groups (Null effect).
* Examples:
* MaskedBenefit\_Example:
  + The intensive program might actually speed up recovery (Benefit). It's likely that patients with the most severe injuries, or those whose recovery was stalling, were the ones channeled into the intensive program (Negative Bias). If the program helped them, it would bring their recovery time back to the average rate. ( Identifies Negative Bias (severe cases/slow recovery select in) and Masking Logic (effectiveness brings them back to average time). )
* MaskedHarm\_Example:
  + The intensive program might actually slow down recovery (Harm), perhaps due to fatigue. The patients who opted for intensive therapy might be the most health-conscious or those with milder injuries who wanted to maximize their recovery (Positive Bias). If the therapy was counterproductive, it would lengthen their recovery time, bringing them down to the average rate. ( Identifies Positive Bias (health-conscious/milder cases select in) and Masking Logic (harm/fatigue brings them down to average time). )
* Neither\_Explained\_Sufficiently\_Example:
  + The study doesn't prove anything because the patients in the intensive group might have been different from the standard group. It's impossible to know the benefits or risks. ( This points out the groups might differ but does not specify how (mechanism) or how this masks an effect (masking logic). )
* Probes:
* Both\_Explained:
  + Thank you. That is a comprehensive interpretation of the intensive therapy results.
  + Thanks, that covers the key possibilities regarding the therapy's benefits and risks.
  + Got it. That's a thorough analysis of the findings.
* MaskedBenefit\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the intensive therapy actually *slowed down* recovery?
  + Thanks. What about the potential risks? Explain how the results might look this way even if the therapy was actually detrimental.
  + Understood. Now, please explain a scenario where the therapy lengthens recovery time, but the study still shows no difference.
* MaskedHarm\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the intensive therapy actually *sped up* recovery?
  + Thanks. What about the potential benefits? Explain how the results might look this way even if the therapy was actually helpful.
  + Understood. Now, please explain a scenario where the therapy shortens recovery time, but the study still shows no difference.
* Neither\_Explained\_Sufficiently:
  + Is the observed 'no difference' sufficient evidence to conclude the intensive therapy has the same impact as the standard protocol? Why or why not?
  + Can you think of a scenario where the intensive therapy *does* change recovery time, but the study still shows no difference?
  + Why might the observed results be misleading about the therapy's actual impact on recovery?
* NotSpecific:
  + Can you be more specific about how that would lead to the observed 'no difference' in recovery time? Please explain the mechanism.
  + That's a bit broad. Can you describe the specific factor and how it might mask the therapy's actual impact?
  + Could you explain the full logic? How exactly does the factor you mentioned lead to the observed 'no difference' result?
* NotDistinct:
  + That relies on the same underlying logic as your previous explanation. Please provide a distinctly different possibility regarding the therapy's benefits or risks.
  + That's too similar to the possibility you already mentioned. Try explaining how the opposite effect could be true.
  + This seems to overlap with an earlier point. Can you describe a different scenario for the therapy's impact?
* NotPlausible:
  + Let's try another angle. Please offer a different interpretation of the results.
  + Okay, please provide a different possibility regarding the therapy's impact.
  + Try to think of a different reason why the study showed no difference.
* NotClear:
  + I didn't follow your reasoning. Could you rephrase your explanation?
  + Can you clarify what you mean? I don't quite understand.
  + I don't quite follow. Could you explain how that would work?
* NotRelevant:
  + Please focus on interpreting the results regarding the intensive therapy program.
  + Your response seems off-topic. Please focus on the benefits and risks of the therapy based on the study.
  + Please address the prompt: how should the administration interpret the 'no difference' finding?

### BDO\_OptionalTraining\_NullEffect\_T1

Stem: A company introduced an optional training program for new employees, intended to improve their career success. However, those who joined the program were no more or less successful than those who didn’t join.

The CEO is now deciding whether to cancel the program.

What would you tell the CEO about the program's actual effectiveness (considering the possibilities that it might be helpful, harmful, or neither) based on this study? Explain your reasoning.

* Mutually Exclusive:
  + Group: training\_program\_zero\_diff
  + With: BDS\_TrainingProgram\_ZeroDiff\_T2
* Scenario:
  + Intervention: Optional training program.
  + Outcome: Career success.
  + Observed Result: Null effect.
* Examples:
* MaskedBenefit\_Example:
  + The program might actually be helpful. If the employees who signed up were mostly those who were already struggling (negative selection), the training might have brought them up to the level of the average employee, making it look like there was no effect.
* MaskedHarm\_Example:
  + The program might actually be harmful, perhaps by wasting time. If the employees who signed up were the most ambitious high-performers (positive selection), they would have done better anyway. The wasted time might have brought their success down to the level of the average employee.
* Neither\_Explained\_Sufficiently\_Example:
  + You can't conclude anything because it wasn't randomized. There was selection bias. ( This identifies the issue (selection bias) but does not provide a specific mechanism or explain how the bias masks a potential effect (Masking Logic). )
* Probes:
* Both\_Explained:
  + Thank you. That is a comprehensive interpretation of the training program results.
  + Thanks, that covers the key possibilities for the training program's impact.
  + Got it. That's a thorough analysis of the findings.
* MaskedBenefit\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the training program actually *harmed* the participants' success?
  + Thanks. What about the other side? Explain how the results might look this way even if the training was actually detrimental.
  + Understood. Now, please explain a scenario where the program actually hurts success, but the study still shows no difference.
* MaskedHarm\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the study show no difference even if the training program actually *helped* the participants' success?
  + Thanks. What about the other side? Explain how the results might look this way even if the training was actually beneficial.
  + Understood. Now, please explain a scenario where the program actually helps success, but the study still shows no difference.
* Neither\_Explained\_Sufficiently:
  + Is the observed 'no difference' sufficient evidence to conclude the training program has no actual impact? Why or why not?
  + Can you think of a scenario where the training program *does* have an impact, but the study still shows no difference?
  + Why might the observed results be misleading about the program's actual effectiveness?
* NotSpecific:
  + Can you be more specific about how that would lead to the observed 'no difference' in success? Please explain the mechanism.
  + That's a bit broad. Can you describe the specific factor and how it might mask the program's actual impact?
  + Could you explain the full logic? How exactly does the factor you mentioned lead to the observed 'no difference' result?
* NotDistinct:
  + That relies on the same underlying logic as your previous explanation. Please provide a distinctly different possibility for the program's actual effect.
  + That's too similar to the possibility you already mentioned. Try explaining how the opposite effect could be true.
  + This seems to overlap with an earlier point. Can you describe a different scenario for the program's impact?
* NotPlausible:
  + Let's try another angle. Please offer a different interpretation of the results.
  + Okay, please provide a different possibility for the program's actual effect.
  + Try to think of a different reason why the study showed no difference.
* NotClear:
  + I didn't follow your reasoning. Could you rephrase your explanation?
  + Can you clarify what you mean? I don't quite understand.
  + I don't quite follow. Could you explain how that would work?
* NotRelevant:
  + Please focus on interpreting the results of the study regarding the training program.
  + Your response seems off-topic. Please focus on the effectiveness of the training program based on the study.
  + Please address the prompt: what is your interpretation of the study results?

### BDO\_RemoteWork\_NullEffect\_T1

Stem: A large tech company allowed its software developers to choose whether to work fully remotely or in the office. Later, an analysis showed that the average productivity metrics were identical for the remote group and the office group.

What is your interpretation of these results regarding how remote work might affect performance, for better or worse? Explain your reasoning.

* Mutually Exclusive:
  + Group: wfh\_job\_satisfaction
  + With: AEG\_WFH\_JobSatisfaction\_T1
* Scenario:
  + Intervention: Working fully remotely (vs. in the office).
  + Outcome: Productivity metrics.
  + Observed Result: No difference in productivity between the two groups (Null effect).
* Examples:
* MaskedBenefit\_Example:
  + Remote work might actually improve productivity (Benefit). Employees who face long commutes or find the office environment distracting might choose remote work (Negative Bias). If working remotely allowed them to focus better, it would raise their productivity up to the average level of the office workers. ( Identifies Negative Bias (long commutes/distracted employees select remote) and Masking Logic (increased focus brings them up to average). )
* MaskedHarm\_Example:
  + Remote work might actually decrease productivity (Harm). The most independent, disciplined, and high-performing employees might be the ones who choose (or are allowed) to work remotely (Positive Bias). If remote work actually hindered collaboration or creativity, it could bring their high productivity down to the average level. ( Identifies Positive Bias (disciplined/high performers select remote) and Masking Logic (reduced collaboration brings them down to average). )
* Neither\_Explained\_Sufficiently\_Example:
  + The results are flawed because employees selected their own location. This means there is self-selection bias, so we can't tell if remote work is better or worse. ( This correctly identifies self-selection bias but does not provide a specific mechanism (e.g., 'disciplined employees') or the masking logic. )
* Probes:
* Both\_Explained:
  + Thank you. That is a comprehensive interpretation of the remote work analysis.
  + Thanks, that covers the key possibilities for how remote work affects performance.
  + Got it. That's a thorough analysis of the productivity data.
* MaskedBenefit\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the analysis show no difference even if working remotely actually *decreased* productivity?
  + Thanks. What about the 'for worse' side? Explain how the results might look this way even if remote work actually hurt productivity.
  + Understood. Now, please explain a scenario where remote work decreases performance, but the analysis still shows no difference.
* MaskedHarm\_Only\_Explained:
  + Okay. Now consider the opposite possibility: how could the analysis show no difference even if working remotely actually *increased* productivity?
  + Thanks. What about the 'for better' side? Explain how the results might look this way even if remote work actually helped productivity.
  + Understood. Now, please explain a scenario where remote work increases performance, but the analysis still shows no difference.
* Neither\_Explained\_Sufficiently:
  + Is the observed 'no difference' sufficient evidence to conclude that work location has no impact on productivity? Why or why not?
  + Can you think of a scenario where remote work *does* affect productivity, but the analysis still shows no difference?
  + Why might the observed results be misleading about the actual impact of remote work?
* NotSpecific:
  + Can you be more specific about how that would lead to the observed 'no difference' in productivity? Please explain the mechanism.
  + That's a bit broad. Can you describe the specific factor and how it might mask the actual impact of remote work?
  + Could you explain the full logic? How exactly does the factor you mentioned lead to the observed 'no difference' result?
* NotDistinct:
  + That relies on the same underlying logic as your previous explanation. Please provide a distinctly different possibility for the impact of remote work.
  + That's too similar to the possibility you already mentioned. Try explaining how the opposite effect could be true.
  + This seems to overlap with an earlier point. Can you describe a different scenario for how remote work affects performance?
* NotPlausible:
  + Let's try another angle. Please offer a different interpretation of the results.
  + Okay, please provide a different possibility regarding the impact of remote work.
  + Try to think of a different reason why the analysis showed no difference.
* NotClear:
  + I didn't follow your reasoning. Could you rephrase your explanation?
  + Can you clarify what you mean? I don't quite understand.
  + I don't quite follow. Could you explain how that would work?
* NotRelevant:
  + Please focus on interpreting the results regarding the impact of remote work on productivity.
  + Your response seems off-topic. Please focus on how remote work affects performance based on the analysis.
  + Please address the prompt: what is your interpretation of these results?

## Schema: BiasDirectionSequential

* Description: Assesses ability to explain a zero-observed difference via selection effects in both directions (benefit vs harm).
* Ability: causal.selection.direction

### BDS\_TrainingProgram\_ZeroDiff\_T2

Stem: A company introduced an optional training program for new employees, intended to improve their career success. However, tehre was no difference in success between those who joined and those who didn’t. Is this good evidence the program has no effect? Why or why not?

* Mutually Exclusive:
  + Group: training\_program\_zero\_diff
  + With: BDO\_OptionalTraining\_NullEffect\_T1
* Scenario:
  + A: joining the optional training program
  + B: career success
* Theme Registry:
  + self\_selection\_motivation
  + prior\_skill\_education
  + manager\_push\_low\_performers
  + workload\_burden\_enrollers
  + cohort\_timing\_hire\_date
* Examples:
* Good:
  + If the program truly helps, lower-performing employees being pushed to enroll could cancel the benefit. (BiasNegative)
  + If the program truly harms, motivated joiners could mask the harm. (BiasPositive)
* NearMisses:
  + Participants learn useful skills; therefore they improve. (NotRelevant)
  + The result is zero, so the program has no effect. (NotRelevant)
* Probes:
* OppositeFromPositive:
  + Now assume the program actually harmed outcomes—what factor could still make both groups look the same?
  + Switch assumptions: if the training reduced success, what factor could cancel that so results show no gap?
* NotDistinct:
  + That overlaps with an earlier reason under the same assumption. Try the opposite direction or a different kind of reason.
* OppositeFromNegative:
  + Now assume the program actually helped—what factor could still make the groups look the same?
  + If the training truly improved success, what factor could offset that so results still look even?
* NotSpecific:
  + Please name a concrete factor and say whether you assume the program helped or harmed.
  + That’s broad—identify a specific factor and note help vs harm.
* NotClear:
  + I didn’t follow—restate your reason and include whether you assume help or harm.
  + Please rephrase more clearly and say help vs harm.
* NotPlausible:
  + Let’s try a more realistic reason that fits the scenario.
* NotRelevant:
  + Please focus on why both groups might look similar even if the program helps or harms.
* MultipleExplanation:
  + One reason at a time, please—send just one.

## Schema: FermiPopulationCity

* Description: Estimate a city's population (orders of magnitude).
* Ability: numeric.fermi

### Fermi\_SF\_Pop\_2020

Stem: Roughly estimate the 2020 population of San Francisco (just give one number).

## Schema: SelectionEffectIdentification

* Description: Assesses the ability to identify and explain selection effects (e.g., confounding by indication, self-selection, survivorship bias) that could explain a surprising or counterintuitive association in an observational context, while adhering to stated premises/constraints.
* Ability: causal.selection\_effects

### SEI\_ClinicalTrial\_Adherence\_T1

Stem: In studying a new drug, it was found that patients who consistently took the drug had better health outcomes than those who frequently missed doses. Assuming neither the drug’s efficacy nor a placebo effect explain this, what's the most likely explanation? (A placebo effect is when the simple act of taking even a sham medical treatment improves outcomes).

* Mutually Exclusive:
  + Group: sei\_adherence\_attrition
  + With: SEI\_WeightLoss\_Attrition\_T1, SEI\_HeartMedication\_Confounding\_T1
* Scenario:
  + A: high adherence (consistently taking the drug)
  + B: better health outcomes
  + Context: Clinical trial for a new drug.
  + Observed Association: High adherence correlated with better outcomes.
  + Constraint 1: Outcome is NOT caused by drug efficacy.
  + Constraint 2: Outcome is NOT caused by placebo effect.
* Theme Registry:
  + healthy\_adherer\_effect\_behavior — compliant people are healthier; adherers follow other health rules; health-conscious behavior; healthy user bias
  + conscientiousness\_confounding — disciplined people; personality factors; responsibility
  + adherence\_confounding\_SES\_stability — more resources; stable living situation; better access to care; social support
  + reverse\_causation\_deterioration — getting sicker leads to missed doses; too sick to take pills; sicker patients miss doses
* Too General: lifestyle, personality, background, the groups were different, confounding variables
* Examples:
* RejectsAssumption:
  + The drug worked for the people who took it consistently. ( Violates the constraint that efficacy is not the cause. )
  + Taking the pills made them feel better due to a placebo. ( Violates the constraint that placebo effect is not the cause. )
* NotSpecific:
  + The adherent people were different from the non-adherent people. ( Too vague—must specify how and why that difference drives outcomes. )
  + It's a confounding variable. ( Names the concept but not the specific factor (e.g., healthy behaviors). )
* Good:
  + People who remember to take pills daily are also more likely to exercise and follow health advice, which improves outcomes. ( Healthy adherer effect. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—send just one.
  + Please provide only one explanation for now.
* Good:
  + Thanks. That is a plausible explanation.
  + Thanks. If you have another distinct explanation, please provide it.
* NotDistinct:
  + Too similar to an earlier idea—try a different reason.
  + This overlaps with a prior answer; please give a different explanation.
* NotSpecific:
  + Please specify the concrete factor and how it affects outcomes.
  + That's broad. How exactly would that lead to the observed result?
* NotClear:
  + I didn’t follow—please put your answer in a different way
  + Sorry, can you clarify what you mean?
* NotPlausible:
  + Let’s try a more realistic reason consistent with the constraints.
  + Let’s try another angle—give a different explanation consistent with the constraints.
* NotRelevant:
  + Please focus on explaining the observed association within the given constraints.
  + Let’s stick to the scenario; address the observed association directly.
* RejectsAssumption:
  + Remember: neither efficacy nor placebo explains this. Offer an explanation consistent with the constraints instead.
  + Please stick to the constraints (no efficacy/placebo); suggest a different selection-based explanation.

### SEI\_CompanyExpansion\_Survivorship\_T1

Stem: An analysis of 10-year-old companies found that those which aggressively expanded in their first year were more profitable than those which grew slowly. Assuming aggressive expansion increases the risk of failure, why might the data suggest it leads to higher profits?

* Mutually Exclusive:
  + With: SEI\_WeightLoss\_Attrition\_T1
* Scenario:
  + A: aggressive early expansion
  + B: higher profits
  + Sample: 10-year-old companies (companies still in business).
  + Observed Association: Aggressive early expansion is correlated with higher profits within this sample.
  + Constraint: Aggressive expansion increases the risk of failure (going out of business).
* Theme Registry:
  + survivorship\_bias\_exclusion\_of\_failures — only looking at the survivors; failed companies are not included; ignoring companies that went bankrupt; data only includes successful firms; missing the failures
* Too General: sampling error, bias in the study, correlation is not causation, they didn't look at all companies, risk vs reward
* Examples:
* RejectsAssumption:
  + Aggressive expansion actually makes companies stronger and more likely to survive. ( Contradicts the premise that aggressive expansion increases failure risk. )
  + You have to take risks to get rewards, so aggressive expansion causes higher profits. ( Argues a direct causal link and ignores the selection mechanism implied by the constraint. )
* NotSpecific:
  + The data is biased. ( True but does not explain how (exclusion of failed firms). )
  + They didn't look at all the data. ( Must specify which data is missing (failed companies) and why that matters. )
* Good:
  + The analysis only includes companies that are still around. Many that expanded aggressively failed and are absent, making the survivors who expanded look more profitable. ( Identifies survivorship bias by exclusion of failures. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please provide only one explanation for now.
* Good:
  + Thank you. That is a plausible explanation.
  + Thanks. If you have another distinct explanation, please provide it.
* NotDistinct:
  + That’s too similar to an earlier reason. Try a different kind of explanation.
  + This overlaps with a prior answer; please give a different mechanism.
* NotSpecific:
  + Please be more specific with your explanation.
  + That's broad. How exactly would that lead to the observed association?
* NotClear:
  + I didn’t follow— could you put your answer a different way?
  + Can you clarify what you mean?
* NotPlausible:
  + Let’s try a more realistic reason consistent with the premise.
  + Let’s try another angle—give a different explanation consistent with the premise.
* NotRelevant:
  + Please focus on explaining the observed association described in the question.
  + Let’s stick to the scenario; address the observed association directly.
* RejectsAssumption:
  + Remember the premise: aggressive expansion increases failure risk. Please offer an explanation consistent with that.
  + Please adhere to the premise (expansion raises failure risk) and provide a selection-based explanation instead.

### SEI\_HeartMedication\_Confounding\_T1

Stem: A large hospital reviews its records and finds that patients who received a new, expensive heart medication have worse survival rates than those who received the old, standard medication. Assuming the new medication is at least as good as the old one, what’s the most plausible reason for this fact?

* Mutually Exclusive:
  + With: SEI\_ClinicalTrial\_Adherence\_T1
* Scenario:
  + A: receiving the new, expensive heart medication
  + B: better survival rates
  + Intervention Group: Patients receiving new, expensive heart medication
  + Comparison Group: Patients receiving old, standard medication
  + Observed Outcome: Intervention group has worse survival rates
  + Constraint: The new medication is at least as good as the old one.
* Theme Registry:
  + confounding\_by\_severity — sicker patients got new drug; worse cases assigned new med; confounding by indication; pre-existing severity
  + last\_resort\_usage — hail mary; failed standard treatment first; used only when near death
  + comorbidity\_frailty\_confounding — other health problems; weaker patients; frailty
* Too General: selection bias, confounding, the groups were different, pre-existing conditions, not randomized, doctor's choice
* Examples:
* RejectsAssumption:
  + The new medication must actually be worse than the old one. ( Violates the constraint that the new medication is at least as good. )
  + The assumption is wrong; the new drug is harmful due to side effects. ( Violates the constraint that the new medication is at least as good. )
* NotSpecific:
  + The patients were not randomized. ( True but does not specify the mechanism causing the difference in outcomes. )
  + There were confounding variables. ( Names the issue but not the specific variable (e.g., severity). )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—could you resend just one?
  + Please provide only one explanation for now.
* Good:
  + Thank you. That is a plausible explanation.
  + Thanks. If you have another distinct explanation, please provide it.
* NotDistinct:
  + That’s too similar to an earlier reason. Try a different kind of explanation.
  + This overlaps with a prior answer; please give a different explanation.
* NotSpecific:
  + Could you be more specific about how the two groups might differ?
  + Can you explain the factor you have in mind?
  + That's broad. How exactly would that lead to the observed result?
* NotClear:
  + I didn’t follow—could you rephrase how that works?
  + Can you clarify the explanation you're giving?
* NotPlausible:
  + Please offer a different kind of explanation for the observation.
  + Let’s try another angle—give a different explanation for the finding.
* NotRelevant:
  + Please focus on explaining the observation described in the prompt.
  + Let’s stick to the scenario; address the observed result directly.
* RejectsAssumption:
  + Remember the constraint that the new medication is at least as good as the old one—please give a reason consistent with that.
  + Please adhere to the given premise (new medication is not worse) and suggest an explanation consistent with it instead.

### SEI\_WeightLoss\_Attrition\_T1

Stem: A 12-week weight loss program reports that the average participant who completed the program lost 30 pounds. Suppose the program is not actually effective but also that the effect was not by chance. How could the reported average weight loss be so high?

* Mutually Exclusive:
  + Group: sei\_adherence\_attrition
  + With: SEI\_ClinicalTrial\_Adherence\_T1, SEI\_CompanyExpansion\_Survivorship\_T1
* Scenario:
  + A: completing the 12-week program
  + B: large weight loss
  + Sample: Participants who completed the 12-week program.
  + Observed Outcome: Average weight loss among completers is high (30 pounds).
  + Constraint: The program is not very effective for most people.
* Theme Registry:
  + attrition\_bias\_dropout\_of\_unsuccessful — people who failed dropped out; only successful people finished; dropouts skew the results; missing data from non-completers; ignoring the dropouts
  + self\_selection\_motivation — highly motivated participants finished; completers were more dedicated
* Too General: bias, sampling issues, not everyone finished, motivation
* Examples:
* RejectsAssumption:
  + The program must actually be highly effective. ( Violates the constraint that the program is not effective. )
  + The program works well, but only for people who stick with it. ( Relies on program efficacy rather than identifying the selection artifact. )
* NotSpecific:
  + Not everyone finished the program. ( Must explain how this leads to the high average (unsuccessful participants dropped out). )
  + They didn't count everyone who started. ( Specify who wasn't counted (dropouts) and why their exclusion biases the average. )
* Good:
  + People who weren't losing weight got discouraged and quit; the reported average counts only those who stayed until the end. ( Identifies attrition bias; unsuccessful dropouts are excluded. )
* Probes:
* MultipleExplanation:
  + One explanation at a time, please—send just one.
  + Please provide only one explanation for now.
* Good:
  + Thanks. That is a plausible explanation.
  + Thanks. If you have another distinct explanation, please provide it.
* NotDistinct:
  + Your answer is too similar to an earlier idea—try a different reason.
  + This overlaps with a prior answer; please give a different explanation.
* NotSpecific:
  + Could you be more specific?
  + That's broad. How exactly would that lead to the observed result?
* NotClear:
  + I didn’t follow—could you rephrase your explanation clearly?
  + Can you clarify what you mean?
* NotPlausible:
  + Let’s try a more realistic reason consistent with the premise.
  + Let’s try another angle—give a different explanation consistent with the premise.
* NotRelevant:
  + Please focus on explaining the observed result described in the question.
  + Let’s stick to the scenario; address the observed result directly.
* RejectsAssumption:
  + Remember the premise: the program is not effective.
  + Please provide an explanation consistent with the premise (not due to program efficacy).