How the public, and scientists, perceive advancement of knowledge from conflicting study results

Derek J. Koehler (Department of Psychology, University of Waterloo)
Gordon Pennycook (Hill/Levene Schools of Business, University of Regina)

Appendix – Scenarios and Measures

Experiment 1 **[W16 UW N=177 3 scenarios x 3 outcomes]**

Educational researchers in a school district in California tested the impact of a new "experiential" (hands-on learning) mathematics curriculum. In three high schools where the new curriculum was introduced, scores on a standardized math test administered at the end of the year were **18% higher** than scores had been the year before.

An independent team of researchers working in a school district in Texas conducted a similar test, and found that standardized math test scores **increased by 15%** [**increased by only 3%**] [**actually decreased by 3%**] compared to scores from the year before the new curriculum was introduced.

[The two teams of researchers identified several potentially important differences in how the new curriculum was implemented and tested in the two studies, such as the amount of teacher training with the new curriculum and the particular standardized test that was administered, as well as differences in student demographics (e.g., family income, cultural background) in the two school districts.]

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the impact of experiential learning on math test scores?

- o We know less
- o We know the same amount
- o We know more

Do these studies advance our knowledge about the impact of experiential learning on math test scores?

- Yes
- No

The body of knowledge built from this area of research seems:

- Very weak
- Weak
- Neither weak nor strong
- Strong
- Very Strong

Should research on this topic receive more, less, or the same continued level of funding?

- Less funding
- The same amount of funding
- More funding

Health researchers in Pennsylvania conducted an experiment in which older adults, aged 50-65, who showed early signs of arthritis, were randomly assigned either to a control (no intervention) condition or to an exercise program that encouraged running or walking at least three times a week. After 3 months, those assigned to the exercise program were **23% less likely** than those in the control group to report that their arthritis symptoms had worsened.

A similar experiment conducted by health researchers in Iowa found that adults in this same age range who were enrolled in an exercise program encouraging running or walking, compared to a control group, were **only 5% less likely** [**actually 5% more likely**] [**19% less likely**] to report that their arthritis symptoms had worsened several months later.

[Health experts who reviewed both experiments noted several differences between the two experiments, such as in how the exercise program was delivered (individual sessions with a trainer versus group classes), the proportion of walkers versus runners in the two studies, the time of day at which arthritis symptoms were reported, and inclusion and exclusion criteria that were used to determine who was eligible for enrollment in the experiments.]

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the exercise program's effect on arthritis symptoms?

- o We know less
- o We know the same amount
- o We know more

Do these studies increase our knowledge about how exercise affects arthritis?

- o Yes
- o No

The body of knowledge built from this area of research seems:

- Very weak
- Weak
- Neither weak nor strong
- Strong
- Very Strong

Should research on this topic receive more, less, or the same continued level of funding?

- Less funding
- The same amount of funding
- More funding

Psychologists at a university in Florida recently completed a study in which they gave people a series of frustrating tasks that were, in fact, impossible to complete, and observed people's nonverbal signs of aggression in response to the tasks, including jaw tightening, fist clenching, and leaning forward. Study participants who reported regularly playing "first-person shooter" videogames were found to display **32% more** such behaviors than were non-players.

A similar study was run independently by psychologists at a university in Arizona. They found that regular players of first-person shooter videogames **actually exhibited 12% fewer** [**exhibited 26% more**] [**only exhibited 12% more**] nonverbal signs of aggression in response to a frustrating task than did non-players.

[The two psychology studies were conducted in somewhat different ways. For instance, the frustrating tasks were not the same in the two studies, nor were the measures of nonverbal aggression identical. The two studies also differed in how often, and how recently, a participant had to play first-person shooter games before they were considered to be a regular player. The pools from which the two studies drew their participants also differed in that one consisted of psychology students receiving extra credit in their introductory psychology course, while the other consisted of students from across the university who were participating for pay.]

When we take the results of these two studies together, do we now know more, less, or the same as we did before about aggression and first-person shooter videogames?

- We know less
- We know the same amount
- We know more

Do these studies increase our knowledge about aggression and first-person shooter videogames?

- Yes
- No

The body of knowledge built from this area of research seems:

- Very weak
- Weak
- Neither weak nor strong
- Strong
- Very Strong

Should research on this topic receive more, less, or the same continued level of funding?

- Less funding
- The same amount of funding
- More funding

Scales

Science / Pseudoscience Beliefs

Trust in Science

Experiment 2 [W16 – Mturk N=411 video game/aggression scenario only 2 agreement x 2 moderators]

Psychologists at a university in Florida recently completed a study in which they gave people a series of frustrating tasks that were, in fact, impossible to complete, and observed people's nonverbal signs of aggression in response to the tasks, including jaw tightening, fist clenching, and leaning forward. Study participants who reported regularly playing "first-person shooter" videogames were found to display **32% more** such behaviors than were non-players. A similar study was run independently by psychologists at a university in Arizona. They found that regular players of first-person shooter videogames **exhibited 26% more [actually exhibited 12% fewe**r] nonverbal signs of aggression in response to a frustrating task than did non-players.

[The two psychology studies were conducted in somewhat different ways. For instance, the frustrating tasks were not the same in the two studies, nor were the measures of nonverbal aggression identical. The two studies also differed in how often, and how recently, a participant had to play first-person shooter games before they were considered to be a regular player. The pools from which the two studies drew their participants also differed in that one consisted of psychology students receiving extra credit in their introductory psychology course, while the other consisted of students from across the university who were participating for pay.]

When we take the results of these two studies together, do we now know more, less, or the same as we did before about aggression and first-person shooter videogames?

- We know less
- We know the same amount
- We know more

Do these studies increase our knowledge about aggression and first-person shooter videogames?

- Yes
- No

The body of knowledge built from this area of research seems:

- Very weak
- Weak
- Neither weak nor strong
- Strong
- Very Strong

Should research on this topic receive more, less, or the same continued level of funding?

- Less funding
- The same amount of funding
- More funding

THEN:
Need for Closure
Trust in Science

Experiment 3 [F16 – Mturk N=941 3 scenarios x 3 outcomes – similar to v2 above but with modified scenario wording and added individual difference measures]

Educational researchers in a school district in California tested the impact of a new "experiential" (hands-on learning) mathematics curriculum. In three high schools where the new curriculum was introduced, scores on a standardized math test administered at the end of the year were **20% higher** than scores had been the year before. An independent team of researchers working in a school district in Texas conducted a similar test, and likewise found that standardized math test scores **increased by 21%** compared to [**were no higher or lower than**]

[**actually decreased by 21%**] scores from the year before the new curriculum was introduced.

An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several potentially important differences in how the new curriculum was implemented and tested in the two studies. For instance, there were differences in the amount of teacher training with the new curriculum and the particular standardized test that was administered, as well as differences in student demographics (e.g., family income, cultural background) in the two school districts.

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the impact of experiential learning on math test scores?

- We know less
- We know the same amount
- We know more

Do these studies advance our knowledge about the impact of experiential learning on math test scores?

- Yes
- No

Health researchers in Pennsylvania conducted an experiment in which older adults, aged 50-65, who showed early signs of arthritis, were randomly assigned either to a control (no intervention) condition or to an exercise program that encouraged running or walking at least three times a week. After 3 months, those assigned to the exercise program were **20% less likely** than those in the control group to report that their arthritis symptoms had worsened. A similar experiment was conducted by health researchers in Iowa and they also found that adults in this same age range who were enrolled in an exercise program encouraging running or walking, compared to a

control group, were **21% less likely** [**no more of less likely**] [**actually 21% more likely**] to report that their arthritis symptoms had worsened several months later.

An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several differences between the two experiments. For instance, there were differences in how the exercise program was delivered (individual sessions with a trainer versus group classes), the proportion of walkers versus runners in the two studies, the time of day at which arthritis symptoms were reported, and inclusion and exclusion criteria that were used to determine who was eligible for enrollment in the experiments.

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the exercise program's effect on arthritis symptoms?

- We know less
- We know the same amount
- We know more

Do these studies increase our knowledge about how exercise affects arthritis?

- Yes
- No

Psychologists at a university in Florida recently completed a study in which they gave people a series of frustrating tasks that were, in fact, impossible to complete, and observed people's nonverbal signs of aggression in response to the tasks, including jaw tightening, fist clenching, and leaning forward. Study participants who reported regularly playing "first-person shooter" videogames were found to display **20% more** such behaviors than were non-players. A similar study was run independently by psychologists at a university in Arizona. They also found that

regular players of first-person shooter videogames **exhibited 21% more** [**no more or fewer**] [**actually exhibited 21% fewer**] nonverbal signs of aggression in response to a frustrating task than did non-players.

An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several differences between the two experiments. For instance, the frustrating tasks were not the same in the two studies, nor were the measures of nonverbal aggression identical. The two studies also differed in how often, and how recently, a participant had to play first-person shooter games before they were considered to be a regular player. The pools from which the two studies drew their participants also differed in that one consisted of psychology students receiving extra credit in their introductory psychology course, while the other consisted of students from across the university who were participating for pay.

When we take the results of these two studies together, do we now know more, less, or the same as we did before about aggression and first-person shooter videogames?

- We know less
- We know the same amount
- We know more

Do these studies increase our knowledge about aggression and first-person shooter videogames?

- Yes
- No

THEN:
Scientific Reasoning Scale
CRT

Experiment 4 [F16 – Mturk v2 N=941 3 scenarios x 3 outcomes]

Two educational research teams, one in California and the other in Texas, recently completed similar experiments on the impact of a new “experiential” (hands-on learning) mathematics curriculum on learning. The studies were conducted independently: Neither research team was aware of the other’s work until after both studies, which were run at about the same time, were completed.

Both studies gave students in three high schools (each) a new curriculum that focused on "experiential" learning with the goal of seeing an improvement in a standardized math test by the end of the year. Because the test is standardized, an improvement would indicate that the "experiential" mathematics curriculum led to more learning than would be expected using more typical mathematics curricula.

In the California study, scores on the standardized math test administered at the end of the year were **20% higher** than scores had been the year before. In the Texas study, scores on the standardized math test administered at the end of the year were **21% higher** [no higher or lower] [**actually 21% lower**] than scores had been the year before.
An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several potentially important differences in how the new curriculum was implemented and tested in the two studies. For instance, there were differences in the amount of teacher training with the new curriculum and the particular standardized test that was administered, as well as differences in student demographics (e.g., family income, cultural background) in the two school districts.

In your opinion, do the results of these two studies, taken together, move us closer toward the truth about the impact of experiential learning on math test scores?

Taken together, relative to where we were before either study was completed, these two studies:

- Move us further from the truth
- Do not move us any closer to the truth
- Move us closer to the truth

When we take the results of these two studies together, do we now know more, less, or the same as we did before the two studies were conducted about the impact of experiential learning on math test scores?

- We know less
- We know the same amount
- We know more

Two health research teams, one in Pennsylvania and the other in Iowa, recently completed similar experiments on the role of exercise in arthritis pain (that is, joint pain in the hands and feet). The studies were conducted independently: Neither research team was aware of the other's work until after both studies, which were run at about the same time, were conducted.

In both studies, older adults (aged 50-65) who showed early signs of arthritis were randomly assigned either to a control (no intervention) condition or to an exercise program that encouraged running or walking at least three times a week. The participants in the study were then contacted after 3 months to see if their arthritis improved.

In the Pennsylvania study, those assigned to the exercise program were **20% more likely** than those in the control group to report that their arthritis symptoms had improved. In the Iowa study, those assigned to the exercise program were 21% more likely [no more or less likely] [**actually** 21% *less* likely] than those in the control group to report that their arthritis symptoms had improved. An independent researcher who reviewed the work found that both studies were

well-executed. However, according to the independent researcher, there were several differences between the two experiments. For instance, there were differences in how the exercise program was delivered (individual sessions with a trainer versus group classes), the proportion of walkers versus runners in the two studies, the time of day at which arthritis symptoms were reported, and inclusion and exclusion criteria that were used to determine who was eligible for enrollment in the experiments.

In your opinion, do the results of these two studies, taken together, move us closer toward the truth about the impact of exercise on arthritis symptoms? Taken together, relative to where we were before either study was completed, these two studies:

- Move us further from the truth
- Do not move us any closer to the truth
- Move us closer to the truth

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the exercise program's effect on arthritis symptoms?

- We know less
- We know the same amount
- We know more

Two university research teams, one in Florida and the other in Arizona, recently completed similar psychological studies testing the relation between aggression and “first-person shooter” video games. The studies were conducted independently: Neither research team was aware of the other’s work until after both studies, which were run at about the same time, were completed.

Both studies gave people a series of frustrating tasks that were, in fact, impossible to complete, and observed people’s nonverbal signs of aggression in response to the tasks. Examples include jaw tightening, fist clenching, and leaning forward.

In the Florida study, participants who reported regularly playing first-person shooter video games displayed **20% more** signs of aggression compared to non-players.
In the Arizona study, participants who reported regularly playing first-person shooter video games displayed **21% more** [**no more (or fewer)**] [**actually displayed 21% *fewer***] signs of aggression compared to non-players.

An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several differences between the two experiments. For instance, the frustrating tasks were not the same in the two studies, nor were the measures of nonverbal aggression identical. The two studies also differed in how often, and how recently, a participant had to play first-person shooter games before they were considered to be a regular player. The pools from which the two studies drew their participants also differed in that one consisted of psychology students receiving extra credit in their introductory psychology course, while the other consisted of students from across the university who were participating for pay.

In your opinion, do the results of these two studies, taken together, move us closer toward the truth about the impact of first-person shooter games on aggression?
Taken together, relative to where we were before either study was completed, these two studies:

- Move us further from the truth
- Do not move us any closer to the truth
- Move us closer to the truth

When we take the results of these two studies together, do we now know more, less, or the same as we did before about aggression and first-person shooter videogames?

- We know less
- We know the same amount
- We know more

Experiment 5 [E5 W17 N=421 2 outcome x 2 sample size]

Researchers were interested in whether people with a particular gene (called Gene X) are more likely to have high blood pressure than people without Gene X.

An **initial study** of 50 randomly-selected adults found evidence supporting an association: The average blood pressure among the 25 adults with Gene X was significantly higher than the average among the 25 adults without Gene X.

A **follow-up study** was then conducted of 100 [50; in this case, number of adults per condition in next paragraph is 25 rather than 50] additional randomly-selected adults.

The follow-up study failed to find an association between Gene X and high blood pressure: The average blood pressure among the 50 adults with Gene X was not significantly higher than the average among the 50 adults without Gene X. [The follow-up study also found an association between Gene X and high blood pressure: The average blood pressure among the 50 adults with Gene X was significantly higher than the average among the 50 adults without Gene X.]

The questions below concern your perception of how the **follow-up study** results change our understanding of the association between Gene X and high blood pressure, relative to what had already been learned from the **initial study**.

In your opinion, relative to where we were after the initial study, do the results of the follow-up study move us closer toward the truth about the association between Gene X and high blood pressure?

- Move us closer to the truth
- Do not move us any closer to the truth
- Move us further from the truth

Compared to what we knew from the results of the initial study, based on the new results from the follow-up study do we now know more, less, or the same as we did before about the association between Gene X and high blood pressure?

- We know more
- We know the same
- We know less

Experiment 6 [E6 F17 AmerAcad N=486 and Mturk N=421 studies; 2 scenarios w/in-S with 2 outcomes btwn-S]

Researchers were interested in whether people with a particular gene (called Gene X) are more likely to have high blood pressure than people without Gene X.

An **initial study** of 50 randomly-selected adults found evidence supporting an association: The average blood pressure among the 25 adults with Gene X was **significantly higher** than the average among the 25 adults without Gene X.

A **follow-up study** was then conducted of 100 additional randomly-selected adults. The follow-up study **failed to find** [**also found**] an association between Gene X and high blood pressure: The average blood pressure among the 50 adults with Gene X was ***not*** significantly higher [was **significantly higher**] than the average among the 50 adults without Gene X.

The questions below concern your perception of how the **follow-up study** results change our understanding of the association between Gene X and high blood pressure, relative to what had already been learned from the **initial study**.

Based on the new results from the follow-up study, do we now know more, less, or the same as we did before about the association between Gene X and high blood pressure?

From the results of the **follow-up study**:

- We know more
- We know the same
- We know less

In your opinion, relative to where we were after the initial study, do the results of the follow-up study move us closer toward the truth about the association between Gene X and high blood pressure?

The results of the **follow-up study**:

- Move us closer to the truth
- Do not move us any closer to the truth
- Move us further from the truth

Two health research teams, one in Pennsylvania and the other in Iowa, recently completed similar experiments on the role of exercise in arthritis pain (that is, joint pain in the hands and feet). The studies were conducted independently: Neither research team was aware of the other's work until after both studies, which were run at about the same time, were conducted.

In both studies, older adults (aged 50-65) who showed early signs of arthritis were randomly assigned either to a control (no intervention) condition or to an exercise program that encouraged running or walking at least three times a week. The participants in the study were then contacted after 3 months to see if their arthritis improved.

In the Pennsylvania study, those assigned to the exercise program were **20% more likely** than those in the control group to report that their arthritis symptoms had improved.

In the Iowa study, those assigned to the exercise program were **actually 21% *less*** likely [were **21% more likely**] than those in the control group to report that their arthritis symptoms had improved.

An independent researcher who reviewed the work found that both studies were well-executed. However, according to the independent researcher, there were several differences between the two experiments. For instance, there were differences in how the exercise program was delivered (individual sessions with a trainer versus group classes), the proportion of walkers versus runners in the two studies, the time of day at which arthritis symptoms were reported, and inclusion and exclusion criteria that were used to determine who was eligible for enrollment in the experiments.

When we take the results of these two studies together, do we now know more, less, or the same as we did before about the impact of exercise on arthritis symptoms?

From the results of **these two studies, taken together**:

- We know more
- We know the same
- We know less

In your opinion, relative to where we were before they were conducted, do the results of these two studies move us closer toward the truth about the impact of exercise on arthritis symptoms?

The results of **these two studies, taken together**:

- Move us closer to the truth
- Do not move us any closer to the truth
- Move us further from the truth

There may be more than one possible explanation for the results of a scientific study.

If a scientific hypothesis is correct, every study that tests it will produce supporting results.

Scientists' knowledge should be called into question when scientific studies produce contradictory results.

[all rated on 7-point strongly disagree to strongly agree scale]