The independence thesis
When individual and social epistemology diverge

Conor Mayo-Wilson    Kevin J.S. Zollman    David Danks

Department of Philosophy
Carnegie Mellon University

Tilburg 2010
Two ways of doing epistemology

What's the relationship between these two?
Two ways of doing epistemology

What's the relationship between these two?
Two ways of doing epistemology

What's the relationship between these two?
Claims of social epistemology

- The drive for credit creates a better allocation of epistemic labor (Kitcher, Goldman, Strevens)
- Dogmatism may make groups more likely to converge on the truth (Feyerabend, Popper, Hull, Kuhn, Zollman)
- Groups of random problem solvers may be better than groups of the best problem solvers (Hong and Page)
- Individually unreliable individuals might pool information to become reliable as a group (Condorcet, Goodin)
- Groups where some information is ignored might outperform groups where information is shared (Elison and Fudenberg, Bala and Goyal, Zollman)
Claims of social epistemology

- The drive for credit creates a better allocation of epistemic labor (Kitcher, Goldman, Strevens)
- Dogmatism may make groups more likely to converge on the truth (Feyerabend, Popper, Hull, Kuhn, Zollman)
- Groups of random problem solvers may be better than groups of the best problem solvers (Hong and Page)
- Individually unreliable individuals might pool information to become reliable as a group (Condorcet, Goodin)
- Groups where some information is ignored might outperform groups where information is shared (Elison and Fudenberg, Bala and Goyal, Zollman)
The drive for credit creates a better allocation of epistemic labor (Kitcher, Goldman, Strevens)

Dogmatism may make groups more likely to converge on the truth (Feyerabend, Popper, Hull, Kuhn, Zollman)

Groups of random problem solvers may be better than groups of the best problem solvers (Hong and Page)

Individually unreliable individuals might pool information to become reliable as a group (Condorcet, Goodin)

Groups where some information is ignored might outperform groups where information is shared (Elison and Fudenberg, Bala and Goyal, Zollman)
Claims of social epistemology

- The drive for credit creates a better allocation of epistemic labor (Kitcher, Goldman, Strevens)
- Dogmatism may make groups more likely to converge on the truth (Feyerabend, Popper, Hull, Kuhn, Zollman)
- Groups of random problem solvers may be better than groups of the best problem solvers (Hong and Page)
- Individually unreliable individuals might pool information to become reliable as a group (Condorcet, Goodin)
- Groups where some information is ignored might outperform groups where information is shared (Elison and Fudenberg, Bala and Goyal, Zollman)
Claims of social epistemology

- The drive for credit creates a better allocation of epistemic labor (Kitcher, Goldman, Strevens)
- Dogmatism may make groups more likely to converge on the truth (Feyerabend, Popper, Hull, Kuhn, Zollman)
- Groups of random problem solvers may be better than groups of the best problem solvers (Hong and Page)
- Individually unreliable individuals might pool information to become reliable as a group (Condorcet, Goodin)
- Groups where some information is ignored might outperform groups where information is shared (Elison and Fudenberg, Bala and Goyal, Zollman)
The independence thesis

1. Epistemically "good" groups might be made up of epistemically "bad" individuals

2. Epistemically "good" individuals might not make up epistemically "good" groups

David Hull

...some of the behavior that appears to be the most improper actually facilitates the manifest goals of science. Mitroff ... remarks that the “problem is how objective knowledge results in science not despite bias and commitment but because of them.” Although objective knowledge through bias and commitment sounds as paradoxical as bombs for peace, I agree that the existence and ultimate rationality of science can be explained in terms of bias, jealousy, and irrationality.
The independence thesis

1. Epistemically “good” groups might be made up of epistemically “bad” individuals
2. Epistemically “good” individuals might not make up epistemically “good” groups

David Hull

...some of the behavior that appears to be the most improper actually facilitates the manifest goals of science. Mitroff ... remarks that the “problem is how objective knowledge results in science not despite bias and commitment but because of them.” Although objective knowledge through bias and commitment sounds as paradoxical as bombs for peace, I agree that the existence and ultimate rationality of science can be explained in terms of bias, jealousy, and irrationality.
The independence thesis

The independence thesis

1. Epistemically “good” groups might be made up of epistemically “bad” individuals
2. Epistemically “good” individuals might not make up epistemically “good” groups

David Hull

...some of the behavior that appears to be the most improper actually facilitates the manifest goals of science. Mitroff ... remarks that the “problem is how objective knowledge results in science not despite bias and commitment but because of them.” Although objective knowledge through bias and commitment sounds as paradoxical as bombs for peace, I agree that the existence and ultimate rationality of science can be explained in terms of bias, jealousy, and irrationality.
The independence thesis

1. Epistemically “good” groups might be made up of epistemically “bad” individuals
2. Epistemically “good” individuals might not make up epistemically “good” groups

Philip Kitcher

... From the community perspective it is likely that sullied scientists will do better than the epistemically pure. This is because a pure community heads toward cognitive uniformity ... By contrast, ... in the sullied community, there are ample opportunities for a division of cognitive labor ... in this way, the sullied community hedges it bets
Outline

1. Independence thesis
   - Two ways of epistemology
   - Independence thesis

2. Our project

3. Individual epistemic propriety
   - Definitions
   - Entailments

4. Group epistemic propriety
   - Definitions
   - Entailments

5. Comparing group and individual standards
   - IC and GIC
   - UC and GUC

6. Conclusion
Our aim

Investigate the independence thesis

- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  
  (We will find that there is more than one way to make this notion precise)

- To prove, mathematically, whether or not these standards conflict
  
  They do . . .
  
  . . . but the degree of conflict depends on which standard one adopts
Our aim

Investigate the independence thesis

- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  - (We will find that there is more than one way to make this notion precise)
- To prove, mathematically, whether or not these standards conflict
  - They do . . .
  - . . . but the degree of conflict depends on which standard one adopts
Our project

Our aim

Investigate the independence thesis

- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  - (We will find that there is more than one way to make this notion precise)
- To prove, mathematically, whether or not these standards conflict
  - They do . . .
  - . . . but the degree of conflict depends on which standard one adopts
Our aim

- Investigate the independence thesis
- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  - (We will find that there is more than one way to make this notion precise)
- To prove, mathematically, whether or not these standards conflict
  - They do . . .
  - . . . but the degree of conflict depends on which standard one adopts
Our project

Our aim

Investigate the independence thesis

- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  - (We will find that there is more than one way to make this notion precise)
- To prove, mathematically, whether or not these standards conflict
  - They do . . .
  - . . . but the degree of conflict depends on which standard one adopts
Our project

Our aim

Investigate the independence thesis

- Make precise what we mean by “individually epistemically good” and “epistemically good group”
  - (We will find that there is more than one way to make this notion precise)
- To prove, mathematically, whether or not these standards conflict
  - They do . . .
  - . . . but the degree of conflict depends on which standard one adopts
The learning problems

- A learner faces a set of “slot machines” each of which has a different distribution over prizes
- The learner can choose to play a slot machine and then observe the outcome
- In some settings, his friends can choose to play the same or different machine and he can observe the outcome of their plays as well
The learning problems

- A learner faces a set of “slot machines” each of which has a different distribution over prizes
- The learner can choose to play a slot machine and then observe the outcome
- In some settings, his friends can choose to play the same or different machine and he can observe the outcome of their plays as well
Our project

The learning problems

- A learner faces a set of “slot machines” each of which has a different distribution over prizes
- The learner can choose to play a slot machine and then observe the outcome
- In some settings, his friends can choose to play the same or different machine and he can observe the outcome of their plays as well
This process is repeated indefinitely

A *method* determines which machine to play given a history of plays and outcomes

There is a tension between exploration and exploitation
Our project

The learning problem

- This process is repeated indefinitely
- A *method* determines which machine to play given a history of plays and outcomes
- There is a tension between exploration and exploitation
The learning problem

- This process is repeated indefinitely
- A *method* determines which machine to play given a history of plays and outcomes
- There is a tension between exploration and exploitation
Why this model?

Medical research

- Set of choices: Each choice represents a potential treatment regime for a disease
- Outcomes: Each outcome represents the all-things-considered effectiveness of the treatment
- States of the world: Different states represent the situation where different treatments are better, all-things-considered
Why this model?

Medical research

- Set of choices: Each choice represents a potential treatment regime for a disease
- Outcomes: Each outcome represents the all-things-considered effectiveness of the treatment
- States of the world: Different states represent the situation where different treatments are better, all-things-considered
Why this model?

Medical research

- Set of choices: Each choice represents a potential treatment regime for a disease
- Outcomes: Each outcome represents the all-things-considered effectiveness of the treatment
- States of the world: Different states represent the situation where different treatments are better, all-things-considered
Why this model?

Medical research

- Set of choices: Each choice represents a potential treatment regime for a disease
- Outcomes: Each outcome represents the all-things-considered effectiveness of the treatment
- States of the world: Different states represent the situation where different treatments are better, all-things-considered
Psychology theory choice: Categories

- Set of choices: Pursue a particular type of explanation for categorization (prototype based, exemplar based, etc.)
- Outcomes: The degree to which that particular application of the explanation-type succeeds in illuminating that phenomenon
- States of the world: The state where one or another theory is better at modeling the phenomenon
Why this model?

Psychology theory choice: Categories

- Set of choices: Pursue a particular type of explanation for categorization (prototype based, exemplar based, etc.)
- Outcomes: The degree to which that particular application of the explanation-type succeeds in illuminating that phenomenon
- States of the world: The state where one or another theory is better at modeling the phenomenon
Psychology theory choice: Categories

- Set of choices: Pursue a particular type of explanation for categorization (prototype based, exemplar based, etc.)
- Outcomes: The degree to which that particular application of the explanation-type succeeds in illuminating that phenomenon
- States of the world: The state where one or another theory is better at modeling the phenomenon
Psychology theory choice: Categories

- Set of choices: Pursue a particular type of explanation for categorization (prototype based, exemplar based, etc.)
- Outcomes: The degree to which that particular application of the explanation-type succeeds in illuminating that phenomenon
- States of the world: The state where one or another theory is better at modeling the phenomenon
Why this model?

Paradigm choice

- Set of choices: The available paradigms – an action of normal science within that paradigm
- Outcomes: The degree to which one succeeds by the paradigm’s own lights
- States of the world: The state where one or another paradigm offers more opportunity for successful normal-science projects
Why this model?

Paradigm choice

- Set of choices: The available paradigms – an action of normal science within that paradigm
- Outcomes: The degree to which one succeeds by the paradigm’s own lights
- States of the world: The state where one or another paradigm offers more opportunity for successful normal-science projects
Why this model?

Paradigm choice

- Set of choices: The available paradigms – an action of normal science within that paradigm
- Outcomes: The degree to which one succeeds by the paradigm’s own lights
- States of the world: The state where one or another paradigm offers more opportunity for successful normal-science projects
Why this model?

Paradigm choice

- Set of choices: The available paradigms – an action of normal science within that paradigm
- Outcomes: The degree to which one succeeds by the paradigm’s own lights
- States of the world: The state where one or another paradigm offers more opportunity for successful normal-science projects
Social interaction

- Individuals can learn from one another by observing their actions and payoffs
- Social interaction is modeled as a fixed, undirected graph
Social interaction

- Individuals can learn from one another by observing their actions and payoffs
- Social interaction is modeled as a fixed, undirected graph
Outline

1. Independence thesis
   - Two ways of epistemology
   - Independence thesis

2. Our project

3. Individual epistemic propriety
   - Definitions
   - Entailments

4. Group epistemic propriety
   - Definitions
   - Entailments

5. Comparing group and individual standards
   - IC and GIC
   - UC and GUC

6. Conclusion
An epistemically “good” individual/group is one that in the long run ends up choosing the best action given the state of the world.
Individual Standards

**Individual consistency**
A method is \textit{individually consistent} if, when implemented by a single individual in isolation, the method converges in to an optimal action.

**Universal consistency**
A method is \textit{universally consistent} if, when implemented by any individual in any social circumstance, the method converges to an optimal action.
Individual Standards

**Individual consistency**
A method is *individually consistent* if, when implemented by a single individual in isolation, the method converges in to an optimal action.

**Universal consistency**
A method is *universally consistent* if, when implemented by any individual in any social circumstance, the method converges to an optimal action.
Individual Standards

Individual consistency
A method is *individually consistent* if, when implemented by a single individual in isolation, the method converges in to an optimal action.

Universal consistency
A method is *universally consistent* if, when implemented by any individual in any social circumstance, the method converges to an optimal action.
Individually and Universally consistent

Since the degenerate graph with only one member is a graph.
Individually and Universally consistent

Trivial example: A strategy which intentionally does something stupid when in a group
Individually and Universally consistent

Non-trivial example:

- **Reinforcement learning** (RL): On each round perform an action in proportion to the *total* past rewards that action has received in the individual’s neighborhood.

- RL is IC
- RL is not UC
Non-trivial example:

- **Reinforcement learning (RL):** On each round perform an action in proportion to the total past rewards that action has received in the individual’s neighborhood.

- RL is IC

- RL is not UC
Individually and Universally consistent

Non-trivial example:
- *Reinforcement learning* (RL): On each round perform an action in proportion to the *total* past rewards that action has received in the individual’s neighborhood.
- RL is IC
- RL is not UC
Outline

1. Independence thesis
   - Two ways of epistemology
   - Independence thesis

2. Our project

3. Individual epistemic propriety
   - Definitions
   - Entailments

4. Group epistemic propriety
   - Definitions
   - Entailments

5. Comparing group and individual standards
   - IC and GIC
   - UC and GUC

6. Conclusion
Group standards

Group individual consistency
Informally, a set of methods $\mathcal{M}$ is *group individually consistent* if a group that implements every strategy in $\mathcal{M}$ will converge when functioning as an isolated group.

Group universal consistency
Informally, a set of methods $\mathcal{M}$ is *group universally consistent* if a group that implements every strategy in $\mathcal{M}$ will converge even when put into an arbitrary community.
**Group standards**

**Group individual consistency**

Informally, a set of methods $M$ is *group individually consistent* if a group that implements every strategy in $M$ will converge when functioning as an isolated group.

**Group universal consistency**

Informally, a set of methods $M$ is *group universally consistent* if a group that implements every strategy in $M$ will converge even when put into an arbitrary community.
Group standards

**Group individual consistency**
Informally, a set of methods $\mathcal{M}$ is *group individually consistent* if a group that implements every strategy in $\mathcal{M}$ will converge *when* functioning as an isolated group.

**Group universal consistency**
Informally, a set of methods $\mathcal{M}$ is *group universally consistent* if a group that implements every strategy in $\mathcal{M}$ will converge *even when* put into an arbitrary community.
Group consistency

Like the individual case, the isolated graph is a graph.
Again, reinforcement learning provides an instructive example.

- Any sized group composed entirely of RL learners is GIC
- But they can each be overwhelmed by manipulating neighbors
Group consistency

Again, reinforcement learning provides an instructive example.

- Any sized group composed entirely of RL learners is GIC
- But they can each be overwhelmed by manipulating neighbors
Group consistency

Again, reinforcement learning provides an instructive example.

- Any sized group composed entirely of RL learners is GIC
- But they can each be overwhelmed by manipulating neighbors
Comparing group and individual standards

Outline

1. Independence thesis
   - Two ways of epistemology
   - Independence thesis

2. Our project

3. Individual epistemic propriety
   - Definitions
   - Entailments

4. Group epistemic propriety
   - Definitions
   - Entailments

5. Comparing group and individual standards
   - IC and GIC
   - UC and GUC

6. Conclusion
Groups vs. Individuals

A group $\mathcal{M}$ – such that each element of $\mathcal{M}$ is IC – is not necessarily GIC.

- Again there is the trivial example: do something stupid when in a group.
- There is a less trivial example.
Groups vs. Individuals

A group $\mathcal{M}$ – such that each element of $\mathcal{M}$ is IC – is not necessarily GIC.

- Again there is the trivial example: do something stupid when in a group.
- There is a less trivial example.
Groups vs. Individuals

A group $\mathcal{M}$ – such that each element of $\mathcal{M}$ is IC – is not necessarily GIC.

- Again there is the trivial example: do something stupid when in a group.
- There is a less trivial example.
**Decreasing epsilon-greedy.** On any given round play take the action with the highest average past payoffs with probability \((1 - \epsilon_i)\) (where \(i\) is the current round). With probability \(\epsilon_i\) take another action with a uniform distribution over the other actions.
If $\epsilon_i$ goes to zero at the right speed (of the order $1/i$) this is IC.

Suppose a version of this strategy where $\epsilon_i = 1/i^k/i$ where $k$ is the total number of observations made thus far.

When alone, $\epsilon_i$ goes to zero at the rate of $1/i$, but in a social setting it goes to zero faster.

As a result, this strategy is IC, but the set containing this single strategy is not GIC.
IC and GIC

- If $\epsilon_i$ goes to zero at the right speed (of the order $1/i$) this is IC.
- Suppose a version of this strategy where $\epsilon_i = 1/i^k$ where $k$ is the total number of observations made thus far.
- When alone, $\epsilon_i$ goes to zero at the rate of $1/i$, but in a social setting it goes to zero faster.
- As a result, this strategy is IC, but the set containing this single strategy is not GIC.
IC and GIC

- If $\epsilon_i$ goes to zero at the right speed (of the order $1/i$) this is IC.
- Suppose a version of this strategy where $\epsilon_i = 1/i^k/i$ where $k$ is the total number of observations made thus far.
- When alone, $\epsilon_i$ goes to zero at the rate of $1/i$, but in a social setting it goes to zero faster.
- As a result, this strategy is IC, but the set containing this single strategy is not GIC.
Comparing group and individual standards

IC and GIC

- If $\epsilon_i$ goes to zero at the right speed (of the order $1/i$) this is IC.

- Suppose a version of this strategy where $\epsilon_i = 1/i^{k/i}$ where $k$ is the total number of observations made thus far.

- When alone, $\epsilon_i$ goes to zero at the rate of $1/i$, but in a social setting it goes to zero faster.

- As a result, this strategy is IC, but the set containing this single strategy is not GIC.
**Delta-epsilon methods** Suppose a strategy which chooses the strategy with the highest average payoff with probability \((1 - \epsilon_i)\) and takes a particular favorite action \(\delta\) with probability \(\epsilon_i\).
IC and GIC

- No version of this strategy is IC
- Some sets of delta-epsilon strategies are GIC
- Let $\mathcal{M}$ be a set of delta-epsilon strategies such that for each action $j$ there is a strategy in $\mathcal{M}$ with $j$ as a favorite action
- $\mathcal{M}$ is GIC
IC and GIC

- No version of this strategy is IC
- Some sets of delta-epsilon strategies are GIC
  - Let $\mathcal{M}$ be a set of delta-epsilon strategies such that for each action $j$ there is a strategy in $\mathcal{M}$ with $j$ as a favorite action
  - $\mathcal{M}$ is GIC
IC and GIC

- No version of this strategy is IC
- Some sets of delta-epsilon strategies are GIC
- Let $\mathcal{M}$ be a set of delta-epsilon strategies such that for each action $j$ there is a strategy in $\mathcal{M}$ with $j$ as a favorite action
  - $\mathcal{M}$ is GIC
IC and GIC

- No version of this strategy is IC
- Some sets of delta-epsilon strategies are GIC
- Let $\mathcal{M}$ be a set of delta-epsilon strategies such that for each action $j$, there is a strategy in $\mathcal{M}$ with $j$ as a favorite action
- $\mathcal{M}$ is GIC
Since UC requires that the method converge in all circumstances, a set containing that method will converge in all social circumstances.
Again, the delta-epsilon collection provides a counter-example.
Our conclusions have implications for how one judges science

- The apparent irrationality of individuals does not imply science as a whole is not rational
  - Our delta-epsilon methods are an example of Popper’s defense of dogmatism
- The apparent rationality of individuals does not imply that science is well functioning
- Our conclusions have normative implications
  - Rationality does not scale up or down.
Conclusions

- Our conclusions have implications for how one judges science
  - The apparent irrationality of individuals does not imply science as a whole is not rational
    - Our delta-epsilon methods are an example of Popper’s defense of dogmatism
  - The apparent rationality of individuals does not imply that science is well functioning
- Our conclusions have normative implications
  - Rationality does not scale up or down
Our conclusions have implications for how one judges science

- The apparent irrationality of individuals does not imply science as a whole is not rational
  - Our delta-epsilon methods are an example of Popper’s defense of dogmatism

- The apparent rationality of individuals does not imply that science is well functioning

Our conclusions have normative implications

- Rationality does not scale up or down
Our conclusions have implications for how one judges science
- The apparent irrationality of individuals does not imply science as a whole is not rational
  - Our delta-epsilon methods are an example of Popper’s defense of dogmatism
- The apparent rationality of individuals does not imply that science is well functioning

Our conclusions have normative implications
- Rationality does not scale up or down.
Conclusions

- Our conclusions have implications for how one judges science
  - The apparent irrationality of individuals does not imply science as a whole is not rational
    - Our delta-epsilon methods are an example of Popper’s defense of dogmatism
  - The apparent rationality of individuals does not imply that science is well functioning
- Our conclusions have normative implications
  - Rationality does not scale up or down.
Conclusions

- Our conclusions have implications for how one judges science
  - The apparent irrationality of individuals does not imply science as a whole is not rational
    - Our delta-epsilon methods are an example of Popper’s defense of dogmatism
  - The apparent rationality of individuals does not imply that science is well functioning
- Our conclusions have normative implications
  - Rationality does not scale up or down.
Further research

- **Strengthen or weaken our standards**
- Consider other models of scientific inquiry
- Some strategies are convergent in particular social settings (myopic Bayesian maximizers)
- Consider more epistemic norms (monotonicity, time to convergence, discounted payoff, etc.)
- Consider game theoretic interaction
Further research

- Strengthen or weaken our standards
- Consider other models of scientific inquiry
  - Some strategies are convergent in particular social settings (myopic Bayesian maximizers)
  - Consider more epistemic norms (monotonicity, time to convergence, discounted payoff, etc.)
  - Consider game theoretic interaction
Further research

- Strengthen or weaken our standards
- Consider other models of scientific inquiry
- Some strategies are convergent in particular social settings (myopic Bayesian maximizers)
  - Consider more epistemic norms (monotonicity, time to convergence, discounted payoff, etc.)
  - Consider game theoretic interaction
Further research

- Strengthen or weaken our standards
- Consider other models of scientific inquiry
- Some strategies are convergent in particular social settings (myopic Bayesian maximizers)
- Consider more epistemic norms (monotonicity, time to convergence, discounted payoff, etc.)
- Consider game theoretic interaction
Further research

- Strengthen or weaken our standards
- Consider other models of scientific inquiry
- Some strategies are convergent in particular social settings (myopic Bayesian maximizers)
- Consider more epistemic norms (monotonicity, time to convergence, discounted payoff, etc.)
- Consider game theoretic interaction