Balancing informational and social goals in active learning

Erica J. Yoon*, Kyle MacDonald*, Mika Asaba,
Hyowon Gweon, & Michael C. Frank

Stanford University

CogSci 2018

Hello, today I'm happy to present our work on balancing informational and social goals in active learning. This work is collaboration with the co-first author Kyle MacDonald, as well as Mika Asaba, Hyo Gweon, and Mike Frank.

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Option 1: New recipe

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Option 1: New recipe

Option 2: Familiar recipe

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Option 1: New recipe

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Option 1: New recipe

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe

Chance to learn ↓

BUT

Risk of failure ↓

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Explore-exploit dilemma

Option 1: New recipe

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe

Chance to learn ↓

BUT

Risk of failure ↓

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Explore-exploit dilemma

Option 1: New recipe Explore

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe

Chance to learn ↓

BUT

Risk of failure ↓

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Explore-exploit dilemma

Option 1: New recipe — Explore

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe — Exploit

Chance to learn ↓

BUT

Risk of failure ↓

Imagine that you are a novice cook, and you want to cook something nice for dinner. And say it comes down to two choices: You can **either** try a completely new recipe you've never tried before, **or** you can just cook up your favorite recipe that is super familiar to you. What are pros and cons? For **the new recipe**, this will be a great opportunity for you to learn to be a great cook, but the risk of failure is pretty high. For the **familiar** recipe, it works the opposite: the risk of failure is low since you're so used to making it, but it'll be one less chance for you to learn to make a potentially awesome new dish.
This is an example of what's known as an **explore-exploit** dilemma, where you have to choose between **exploring** a new domain to gain new knowledge, versus **exploiting** your previous knowledge to ensure some immediate positive return.

Explore-exploit dilemma

Option 1: New recipe — Explore

Chance to learn ↑

BUT

Risk of failure ↑

Option 2: Familiar recipe — Exploit

Chance to learn ↓

BUT

Risk of failure ↓

Now, imagine that you are not just cooking for yourself but instead, you're cooking for your date. How would that affect your decision-making process?

Now you're probably re-thinking your pros and cons with respect to appearing awesome to your date — So for the **new recipe**, it'll be a great chance to learn AND to show off if you do succeed, but then you risk being totally embarrassed if you utterly fail at your first attempt at this new dish. On the other hand, choosing the **familiar** recipe will sacrifice one opportunity to learn, but it will probably at least guarantee that you won't fail and embarrass yourself in front of your date. Like this, **social factors** can influence our decision-making process when we're deciding to explore or to exploit.

Explore-exploit dilemma

Option 1: New recipe — Explore

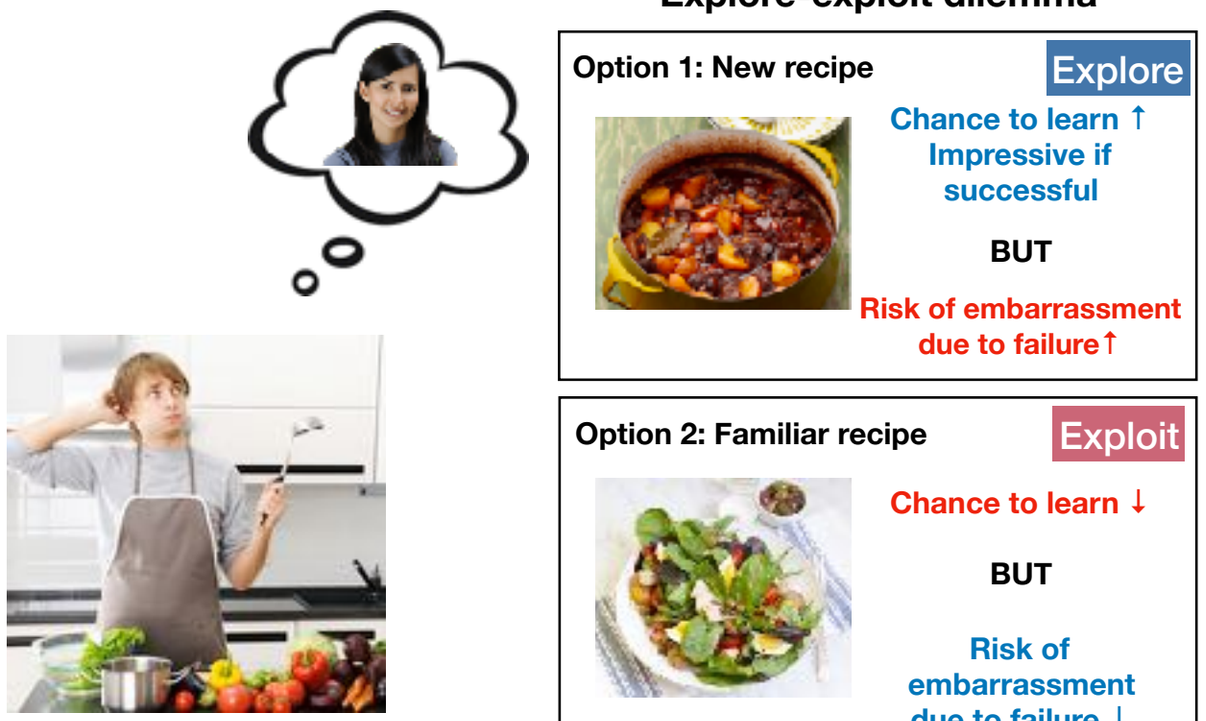
Chance to learn ↑
Impressive if successful

BUT

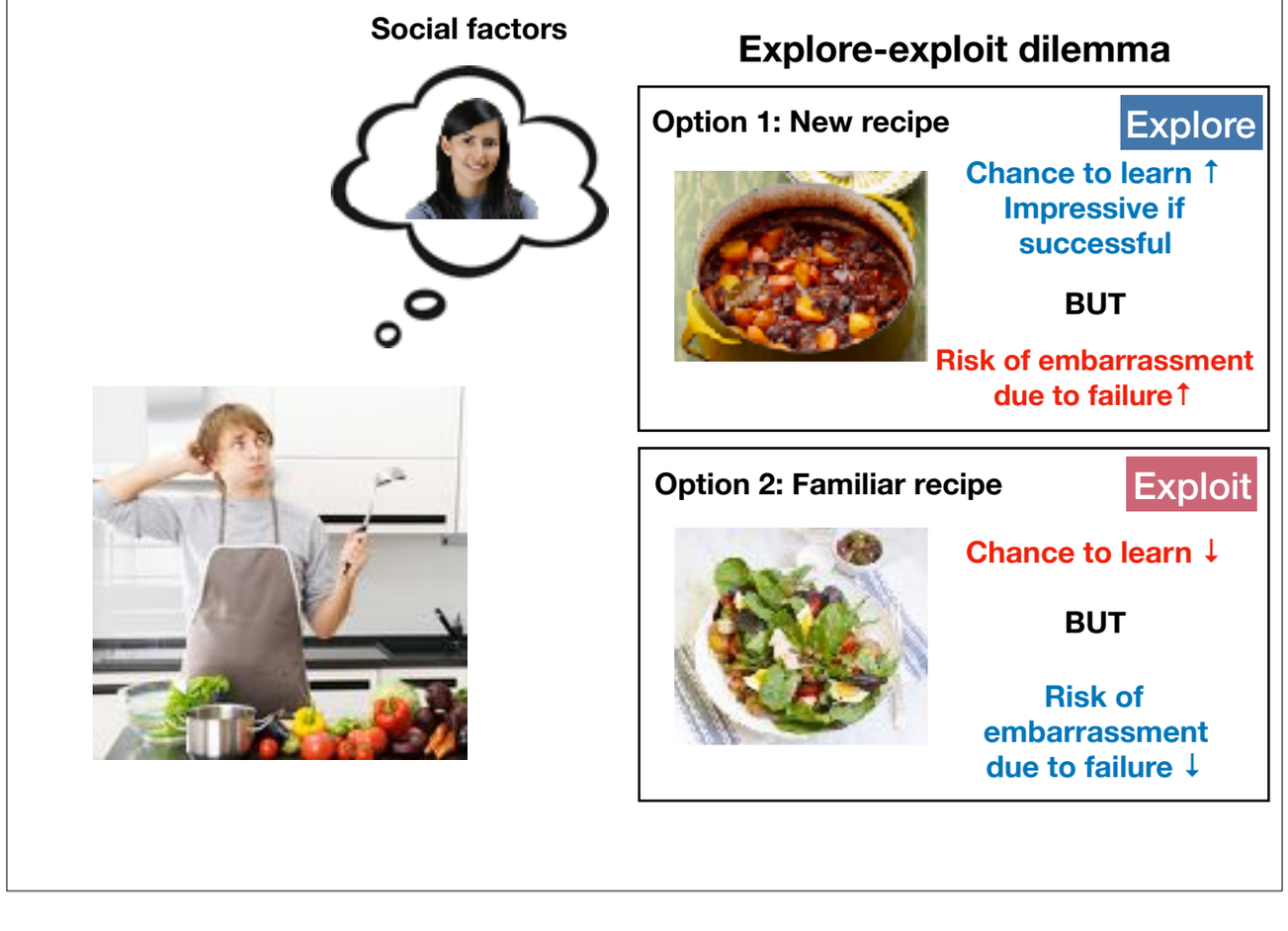
Risk of embarrassment due to failure↑

Option 2: Familiar recipe — Exploit

Now you're probably re-thinking your pros and cons with respect to appearing awesome to your date — So for the **new recipe**, it'll be a great chance to learn AND to show off if you do succeed, but then you risk being totally embarrassed if you utterly fail at your first attempt at this new dish. On the other hand, choosing the **familiar** recipe will sacrifice one opportunity to learn, but it will probably at least guarantee that you won't fail and embarrass yourself in front of your date. Like this, **social factors** can influence our decision-making process when we're deciding to explore or to exploit.



Now you're probably re-thinking your pros and cons with respect to appearing awesome to your date — So for the **new recipe**, it'll be a great chance to learn AND to show off if you do succeed, but then you risk being totally embarrassed if you utterly fail at your first attempt at this new dish. On the other hand, choosing the **familiar** recipe will sacrifice one opportunity to learn, but it will probably at least guarantee that you won't fail and embarrass yourself in front of your date. Like this, **social factors** can influence our decision-making process when we're deciding to explore or to exploit.



Now you're probably re-thinking your pros and cons with respect to appearing awesome to your date — So for the **new recipe**, it'll be a great chance to learn AND to show off if you do succeed, but then you risk being totally embarrassed if you utterly fail at your first attempt at this new dish. On the other hand, choosing the **familiar** recipe will sacrifice one opportunity to learn, but it will probably at least guarantee that you won't fail and embarrass yourself in front of your date. Like this, **social factors** can influence our decision-making process when we're deciding to explore or to exploit.

In the current work, we start with this question of: how do social factors influence our goals in an active learning context?

Outline

- Active learning (+ social factors)
- Model of goal tradeoffs in active learning
  - Paradigm
  - Model structure
- Experimental task
  - Behavioral results
  - Model predictions

So in this talk, I will first talk about what we mean by active learning and social factors that may affect its process,
And then I'll present our model of tradeoffs of informational and social goals that we might consider in an active learning context,
And finally I will discuss an experimental task we designed to test our predictions from the model.

Outline

- **Active learning (+ social factors)**
- Model of goal tradeoffs in active learning
  - Paradigm
  - Model structure
- Experimental task
  - Behavioral results
  - Model predictions

So first, what is active learning?

Active learning

Active learning occurs when learners have control over the sequence of information during learning. You may choose which **dish** to cook to learn more about cooking, or, in another familiar example, you may decide to **raise** your hand to ask a question after a talk to learn more information about it.
Previous research in cognitive science has tried to capture active learning in a quantifiable context, and so have mostly focused on causal learning, where, for example, people choose to press buttons on a toy, one by one, to learn the effect of each button.

Active learning

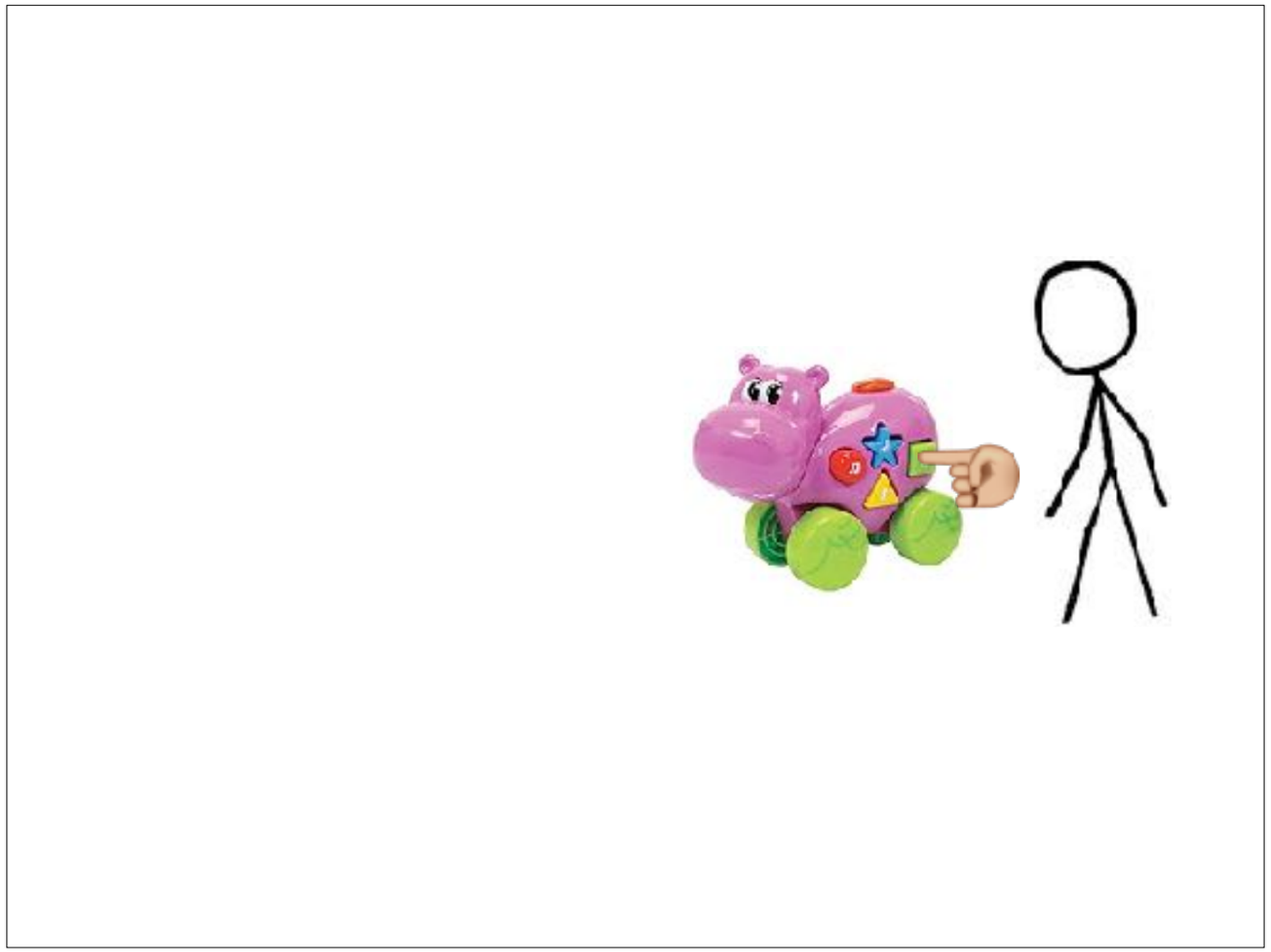
Active learning occurs when learners have control over the sequence of information during learning. You may choose which **dish** to cook to learn more about cooking, or, in another familiar example, you may decide to **raise** your hand to ask a question after a talk to learn more information about it.
Previous research in cognitive science has tried to capture active learning in a quantifiable context, and so have mostly focused on causal learning, where, for example, people choose to press buttons on a toy, one by one, to learn the effect of each button.

Active learning

Active learning occurs when learners have control over the sequence of information during learning. You may choose which **dish** to cook to learn more about cooking, or, in another familiar example, you may decide to **raise** your hand to ask a question after a talk to learn more information about it.
Previous research in cognitive science has tried to capture active learning in a quantifiable context, and so have mostly focused on causal learning, where, for example, people choose to press buttons on a toy, one by one, to learn the effect of each button.

Active learning

Active learning occurs when learners have control over the sequence of information during learning. You may choose which **dish** to cook to learn more about cooking, or, in another familiar example, you may decide to **raise** your hand to ask a question after a talk to learn more information about it.
Previous research in cognitive science has tried to capture active learning in a quantifiable context, and so have mostly focused on causal learning, where, for example, people choose to press buttons on a toy, one by one, to learn the effect of each button.



Here, we also use a causal learning context to see how social factors may affect active learning process. Previous work has looked at how other people's actions may affect a learner's inferences and action, but much less work has examined the effects of social context on active learning. in other words, your choice for whether you should press the buttons on a machine or make a new **dish** may not just depend on your want to learn, but also whether it'll make you look good to **others**.
We began our work with this question of how social goals might influence the learner's actions — more **specifically**, how do people integrate …

Here, we also use a causal learning context to see how social factors may affect active learning process. Previous work has looked at how other people's actions may affect a learner's inferences and action, but much less work has examined the effects of social context on active learning. in other words, your choice for whether you should press the buttons on a machine or make a new **dish** may not just depend on your want to learn, but also whether it'll make you look good to **others**.
We began our work with this question of how social goals might influence the learner's actions — more **specifically**, how do people integrate …

Here, we also use a causal learning context to see how social factors may affect active learning process. Previous work has looked at how other people's actions may affect a learner's inferences and action, but much less work has examined the effects of social context on active learning. in other words, your choice for whether you should press the buttons on a machine or make a new **dish** may not just depend on your want to learn, but also whether it'll make you look good to **others**.
We began our work with this question of how social goals might influence the learner's actions — more **specifically**, how do people integrate …

How do people integrate **social goals** with **informational goals** when deciding what to do in an active learning context?

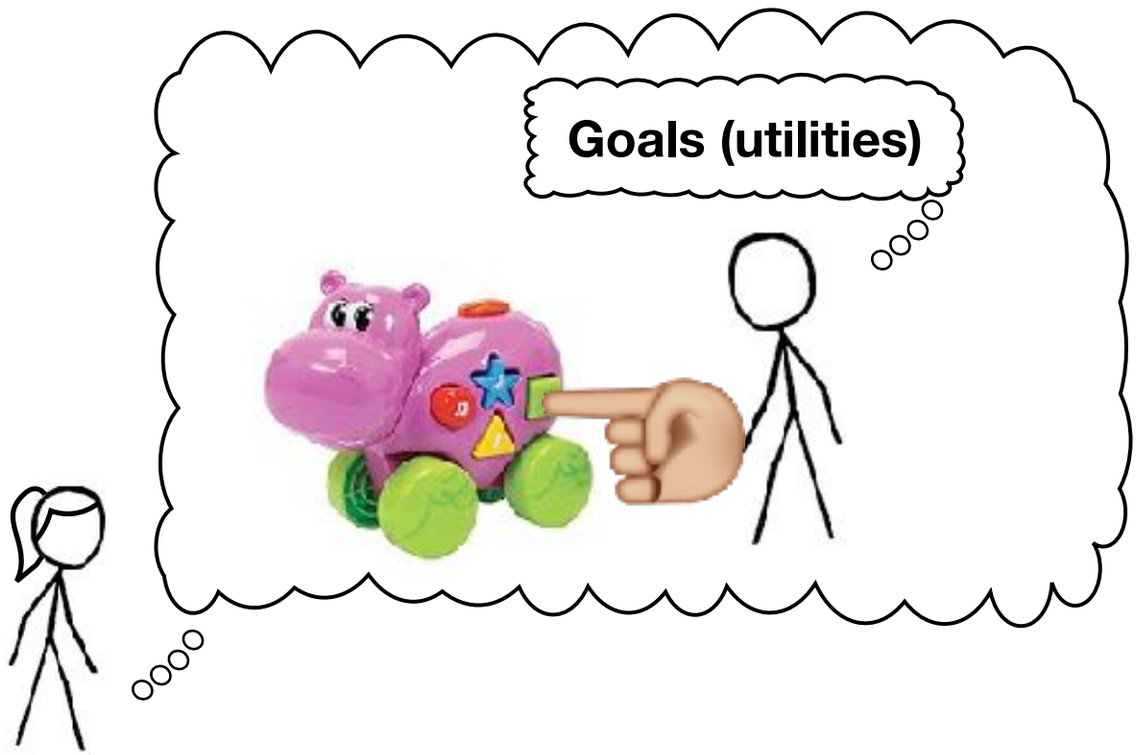
Here, we also use a causal learning context to see how social factors may affect active learning process. Previous work has looked at how other people's actions may affect a learner's inferences and action, but much less work has examined the effects of social context on active learning. in other words, your choice for whether you should press the buttons on a machine or make a new **dish** may not just depend on your want to learn, but also whether it'll make you look good to **others**.
We began our work with this question of how social goals might influence the learner's actions — more **specifically**, how do people integrate …

Outline

- Active learning (+ social factors)
- **Model of goal tradeoffs in active learning**
  - Paradigm
  - Model structure
- Experimental task
  - Behavioral results
  - Model predictions

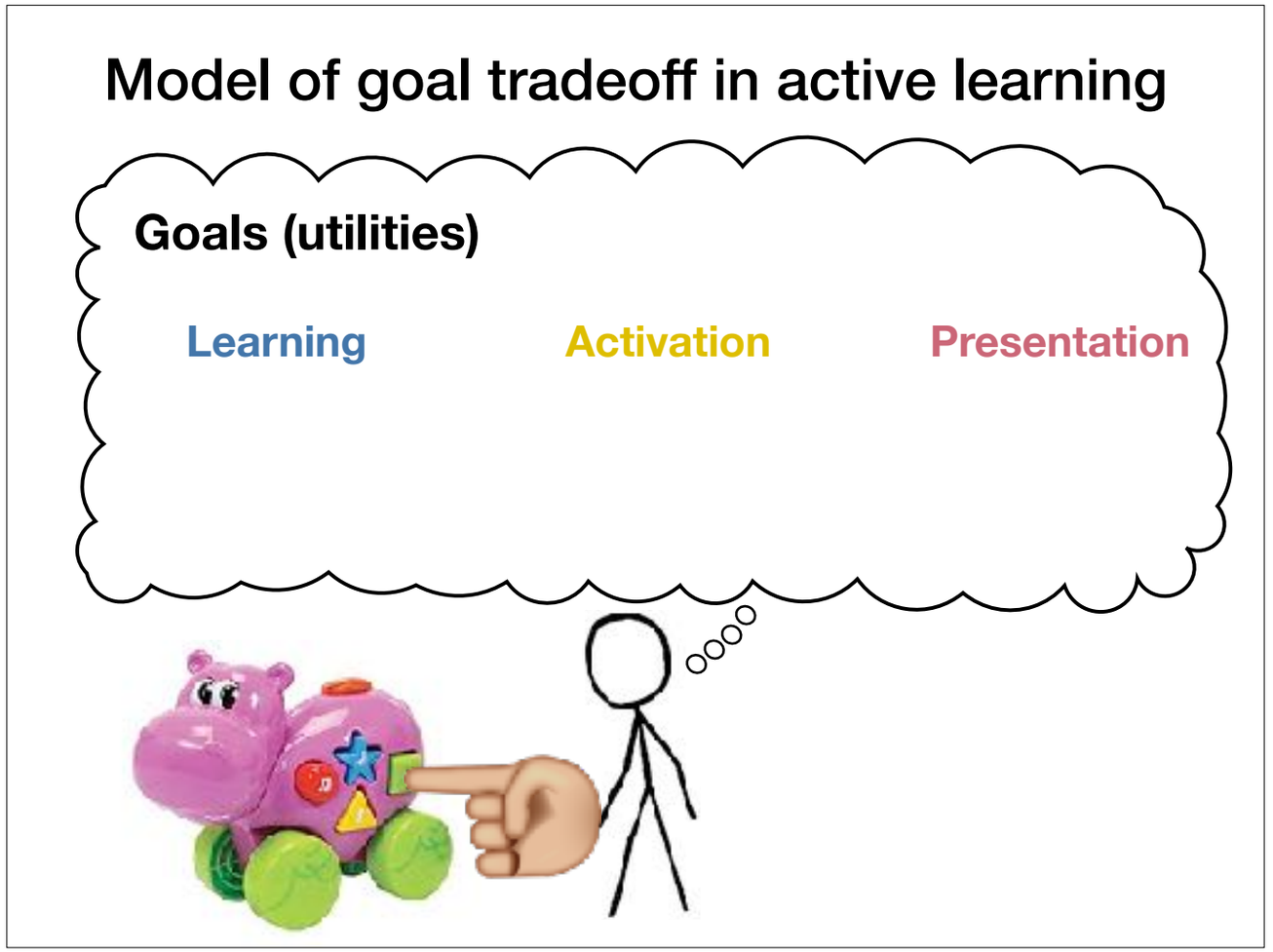
To start to answer this question, we model a learner who considers a mixture of goals, and their tradeoffs in a simple active learning context.

Model of goal tradeoff in active learning

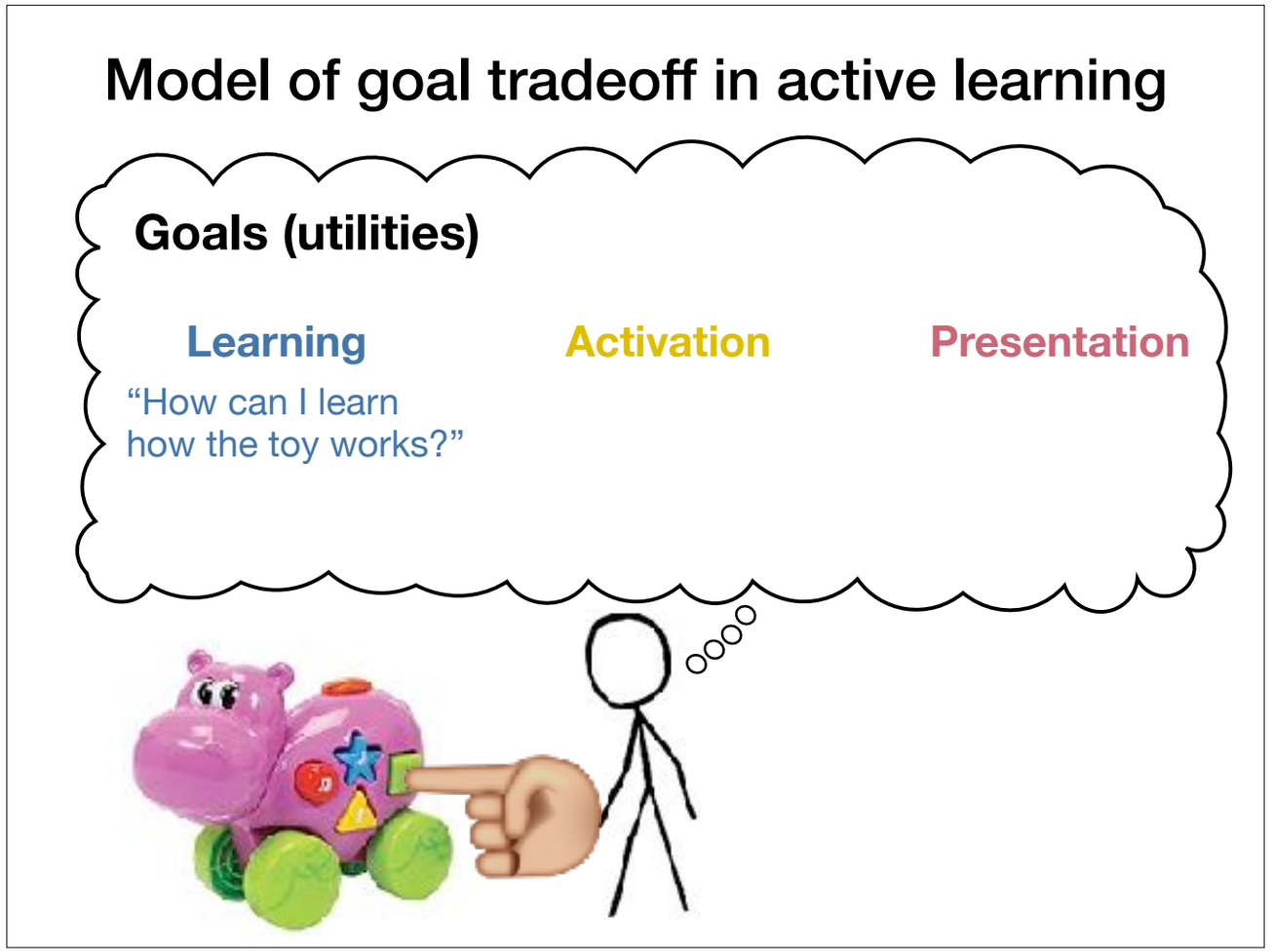


Goodman & Frank, 2016; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016

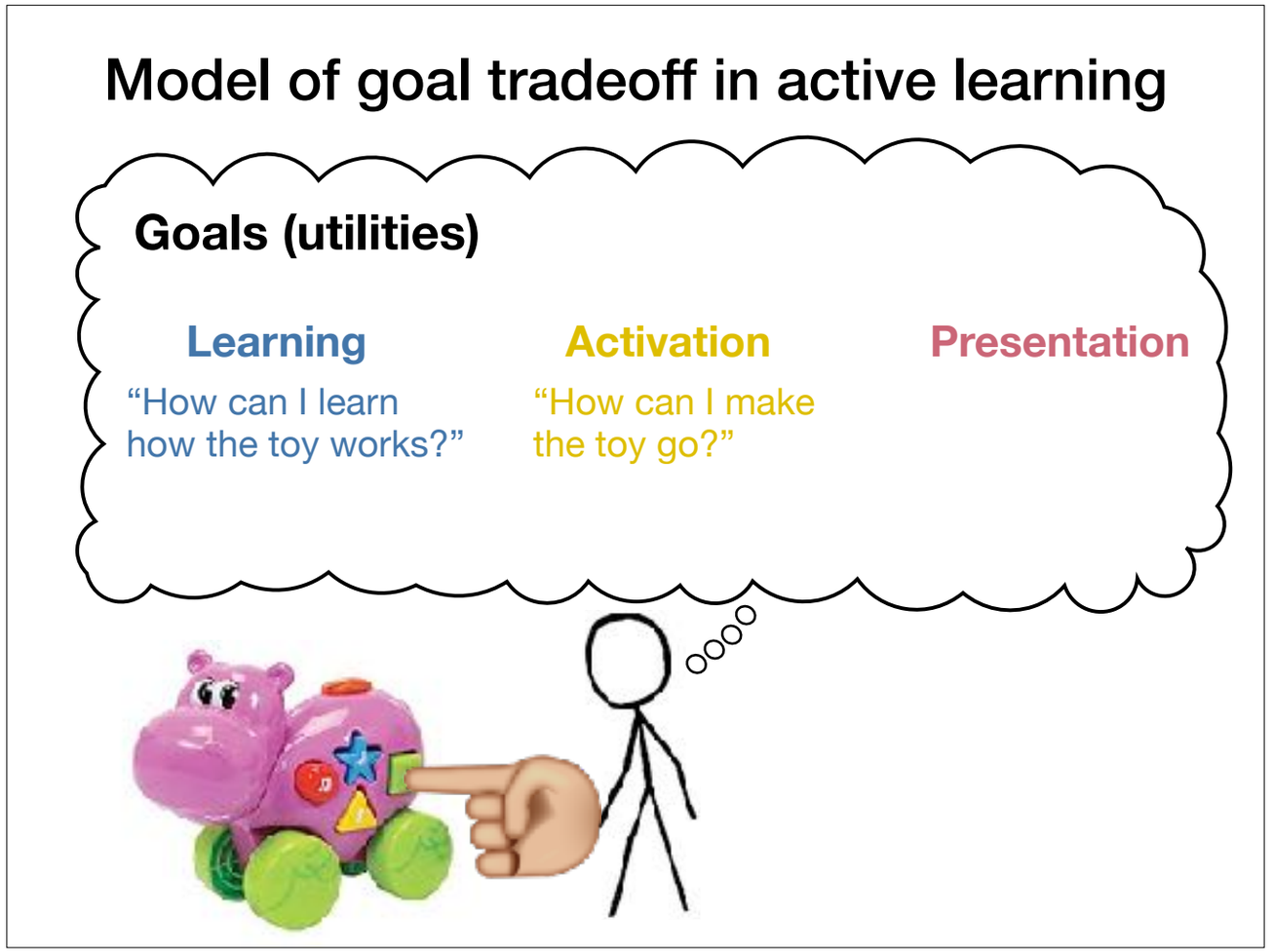
To build this model, we take the key assumption underlying recent Bayesian models of human social cognition, that people expect others to act approximately optimally given a utility function that represents their goal.



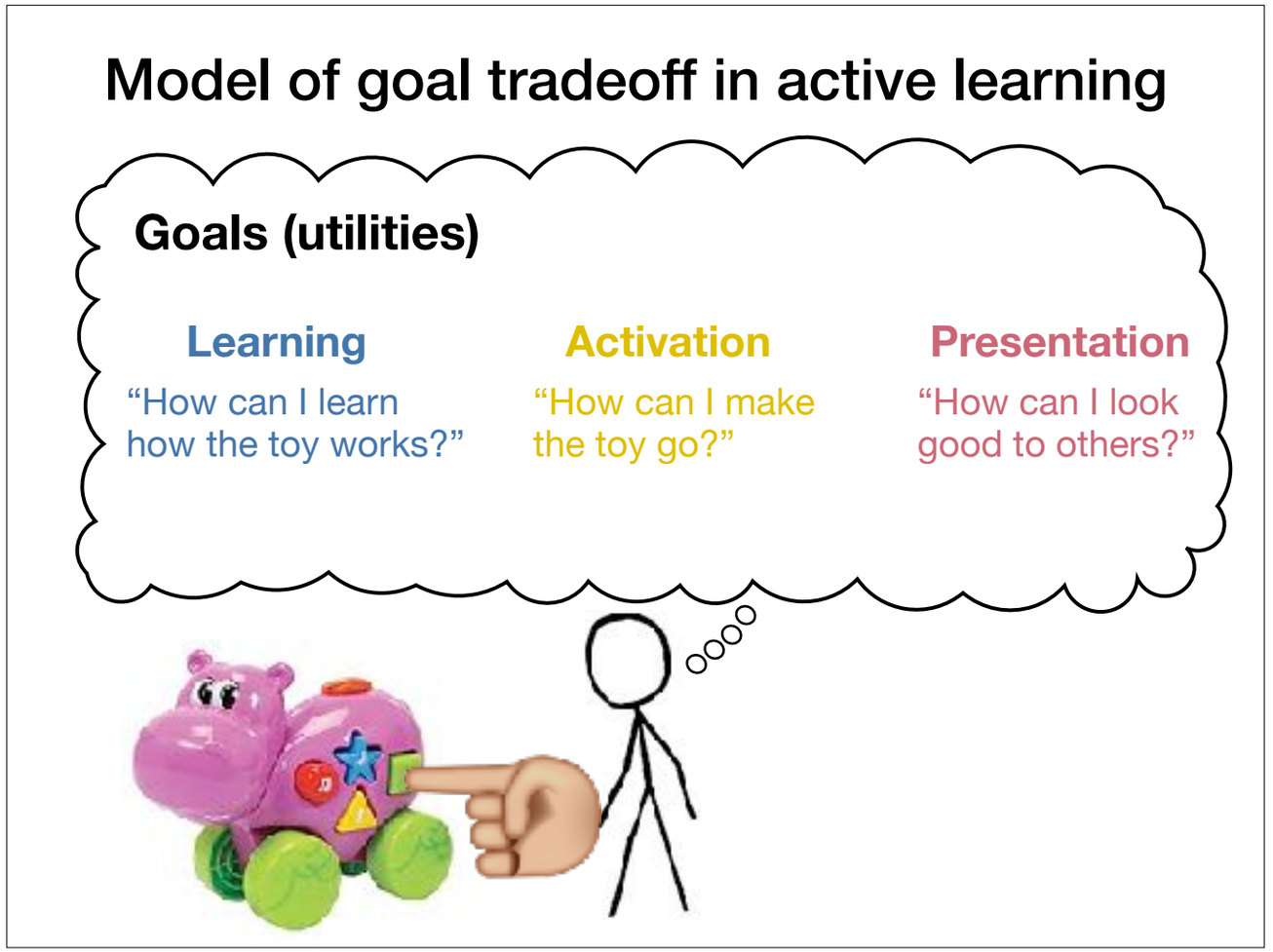
Our model adopts this utility-theoretic approach and assumes an agent who reasons about three different goals: a **learning** goal, which has to do with how he can learn how a given toy works, an **activation** goal, about how he can make the toy go, and a **presentation** goal, to do with how he can look good to other people. We use a utility function that represents a **weighted** combination of these goals that we have implemented in our previous work in a different domain, looking at polite speech production as a tradeoff between goals to be informative versus goals to appear kind to other people. So we applied this same idea of the tradeoff between informational and social goals.



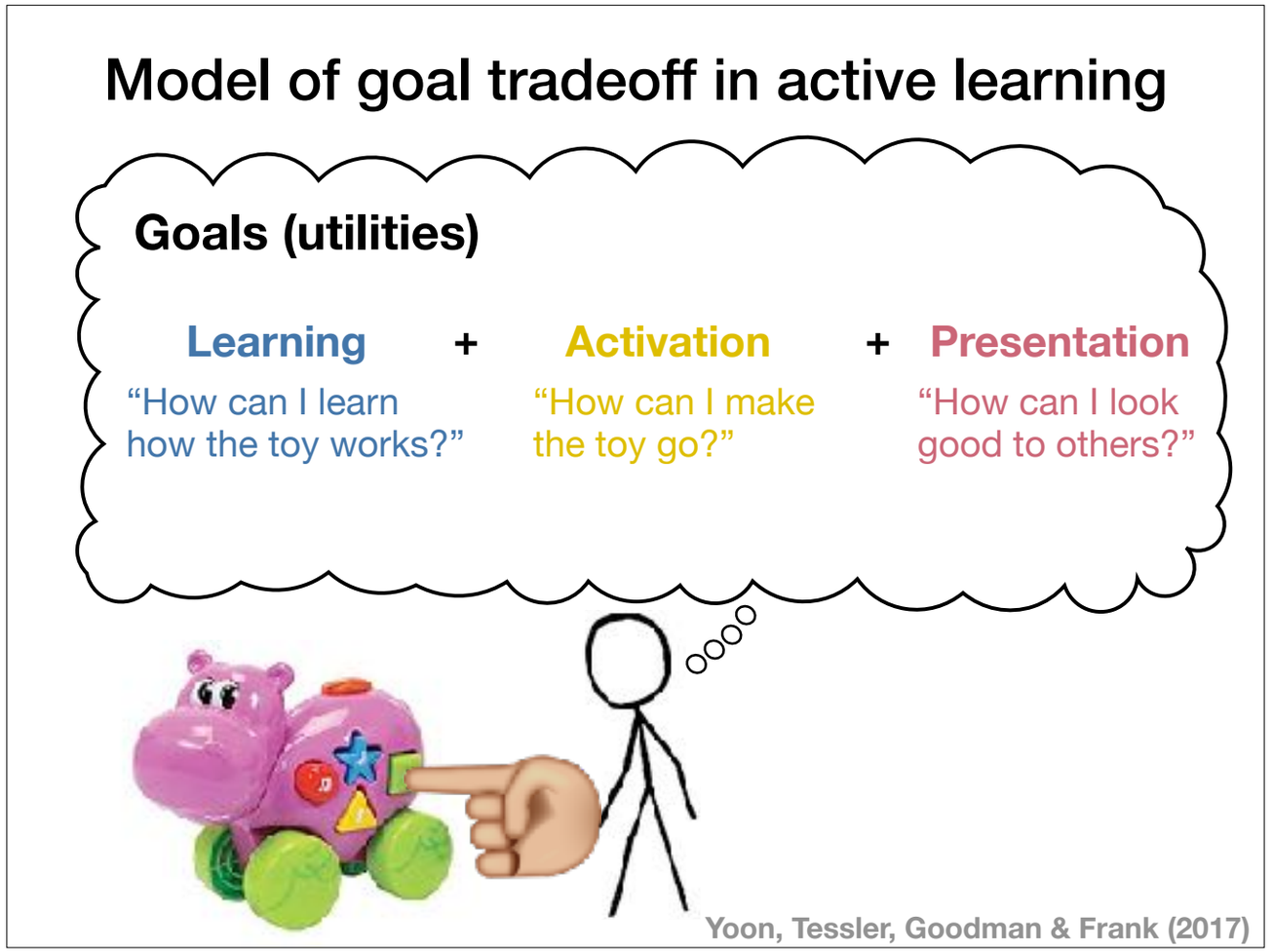
Our model adopts this utility-theoretic approach and assumes an agent who reasons about three different goals: a **learning** goal, which has to do with how he can learn how a given toy works, an **activation** goal, about how he can make the toy go, and a **presentation** goal, to do with how he can look good to other people. We use a utility function that represents a **weighted** combination of these goals that we have implemented in our previous work in a different domain, looking at polite speech production as a tradeoff between goals to be informative versus goals to appear kind to other people. So we applied this same idea of the tradeoff between informational and social goals.



Our model adopts this utility-theoretic approach and assumes an agent who reasons about three different goals: a **learning** goal, which has to do with how he can learn how a given toy works, an **activation** goal, about how he can make the toy go, and a **presentation** goal, to do with how he can look good to other people. We use a utility function that represents a **weighted** combination of these goals that we have implemented in our previous work in a different domain, looking at polite speech production as a tradeoff between goals to be informative versus goals to appear kind to other people. So we applied this same idea of the tradeoff between informational and social goals.

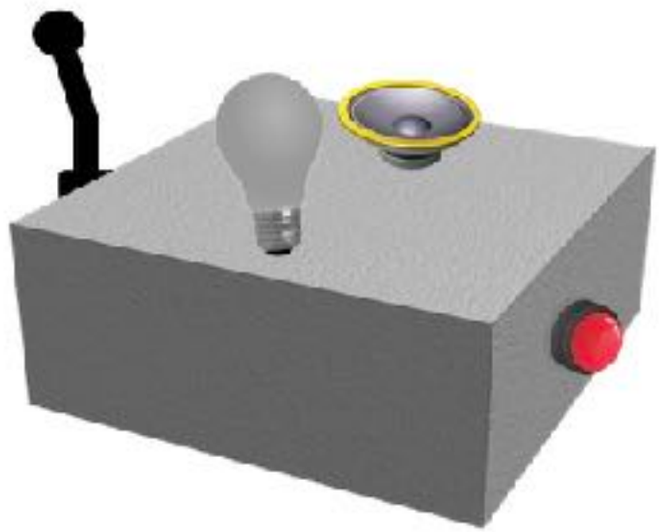


Our model adopts this utility-theoretic approach and assumes an agent who reasons about three different goals: a **learning** goal, which has to do with how he can learn how a given toy works, an **activation** goal, about how he can make the toy go, and a **presentation** goal, to do with how he can look good to other people. We use a utility function that represents a **weighted** combination of these goals that we have implemented in our previous work in a different domain, looking at polite speech production as a tradeoff between goals to be informative versus goals to appear kind to other people. So we applied this same idea of the tradeoff between informational and social goals.



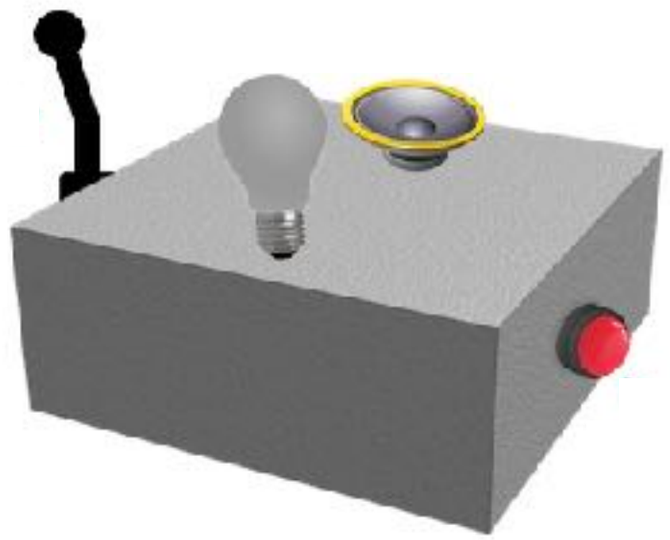
Our model adopts this utility-theoretic approach and assumes an agent who reasons about three different goals: a **learning** goal, which has to do with how he can learn how a given toy works, an **activation** goal, about how he can make the toy go, and a **presentation** goal, to do with how he can look good to other people. We use a utility function that represents a **weighted** combination of these goals that we have implemented in our previous work in a different domain, looking at polite speech production as a tradeoff between goals to be informative versus goals to appear kind to other people. So we applied this same idea of the tradeoff between informational and social goals.

Simple active learning paradigm



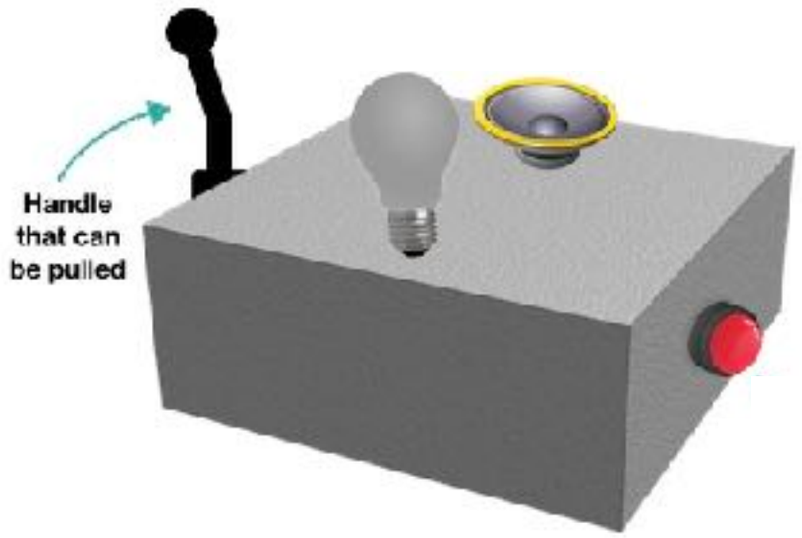
Our model is instantiated in a simple causal learning task, where there are different actions associated with different utilities. We present a toy with an ambiguous causal mechanism that looks like this.

Simple active learning paradigm



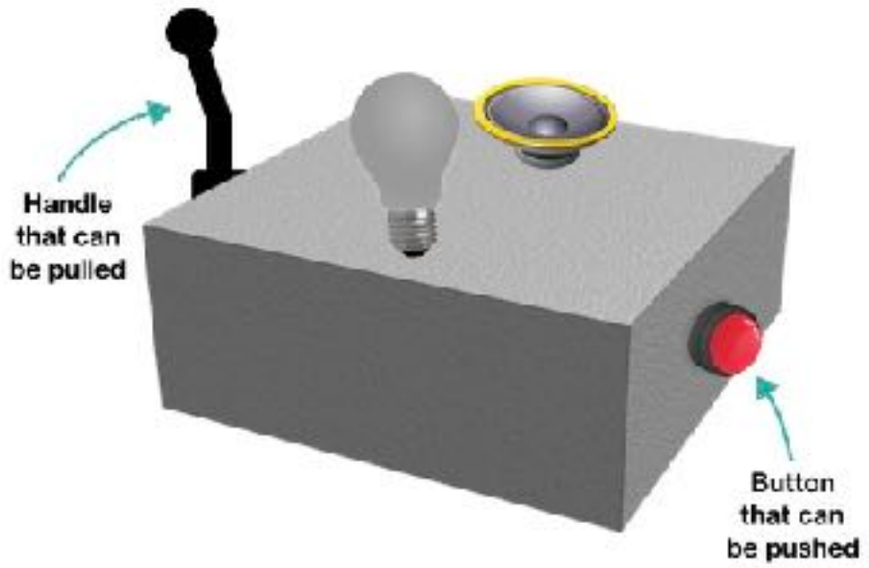
For this toy, there are two things that can be acted upon: a **handle** that can be pulled, and a **button** that can be pushed. These things are somehow linked to these **two** other things on top of the toy: a light that can turn on and a speaker where the sound can play. However, this toy is **missing** its label so right now it's impossible to know how this particular toy works.

Simple active learning paradigm



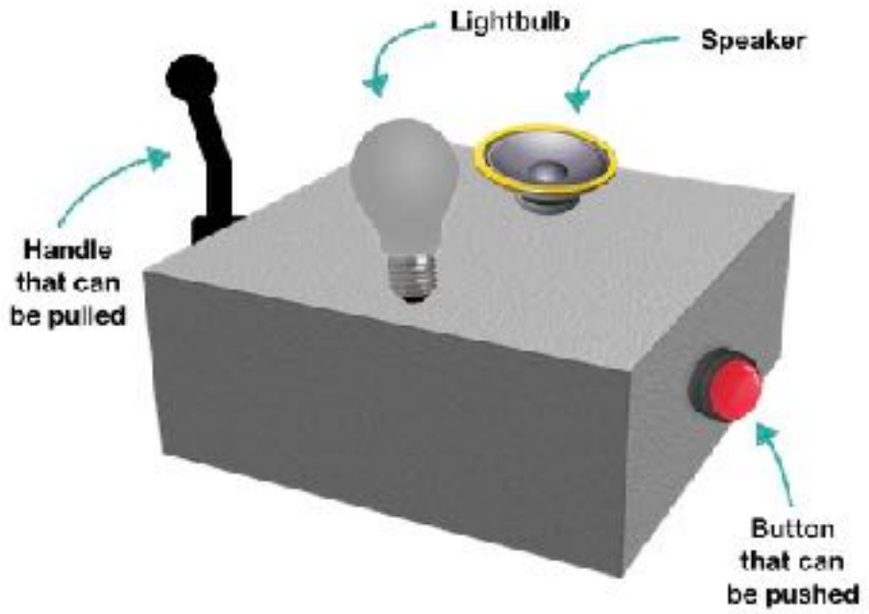
For this toy, there are two things that can be acted upon: a **handle** that can be pulled, and a **button** that can be pushed. These things are somehow linked to these **two** other things on top of the toy: a light that can turn on and a speaker where the sound can play. However, this toy is **missing** its label so right now it's impossible to know how this particular toy works.

Simple active learning paradigm



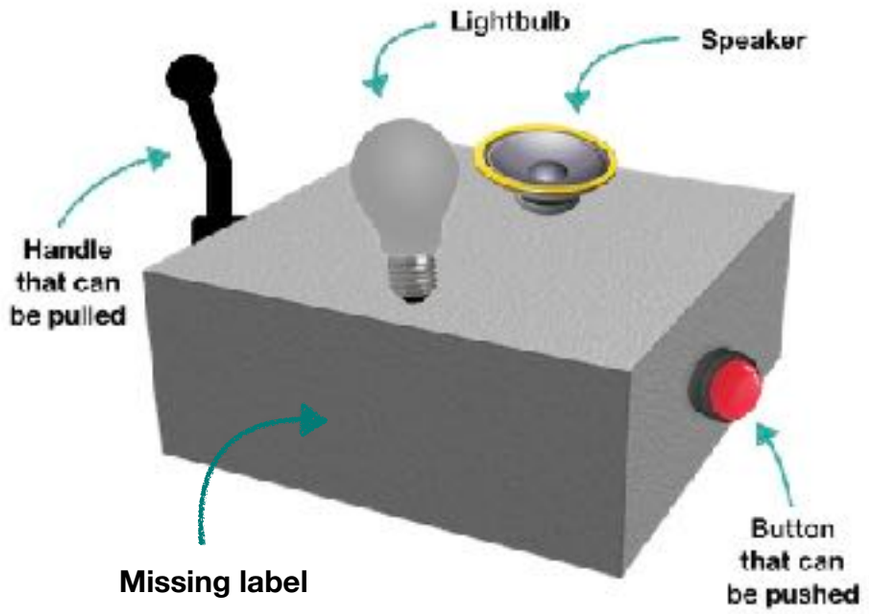
For this toy, there are two things that can be acted upon: a **handle** that can be pulled, and a **button** that can be pushed. These things are somehow linked to these **two** other things on top of the toy: a light that can turn on and a speaker where the sound can play. However, this toy is **missing** its label so right now it's impossible to know how this particular toy works.

Simple active learning paradigm



For this toy, there are two things that can be acted upon: a **handle** that can be pulled, and a **button** that can be pushed. These things are somehow linked to these **two** other things on top of the toy: a light that can turn on and a speaker where the sound can play. However, this toy is **missing** its label so right now it's impossible to know how this particular toy works.

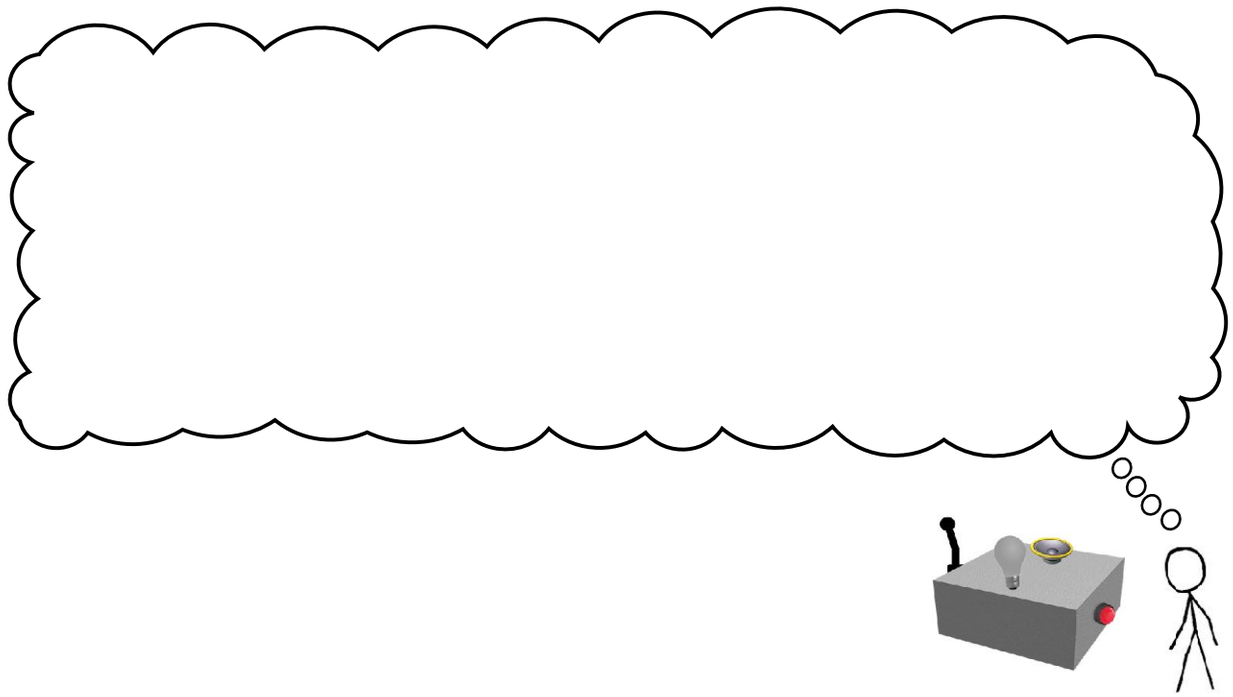
Simple active learning paradigm



For this toy, there are two things that can be acted upon: a **handle** that can be pulled, and a **button** that can be pushed. These things are somehow linked to these **two** other things on top of the toy: a light that can turn on and a speaker where the sound can play. However, this toy is **missing** its label so right now it's impossible to know how this particular toy works.

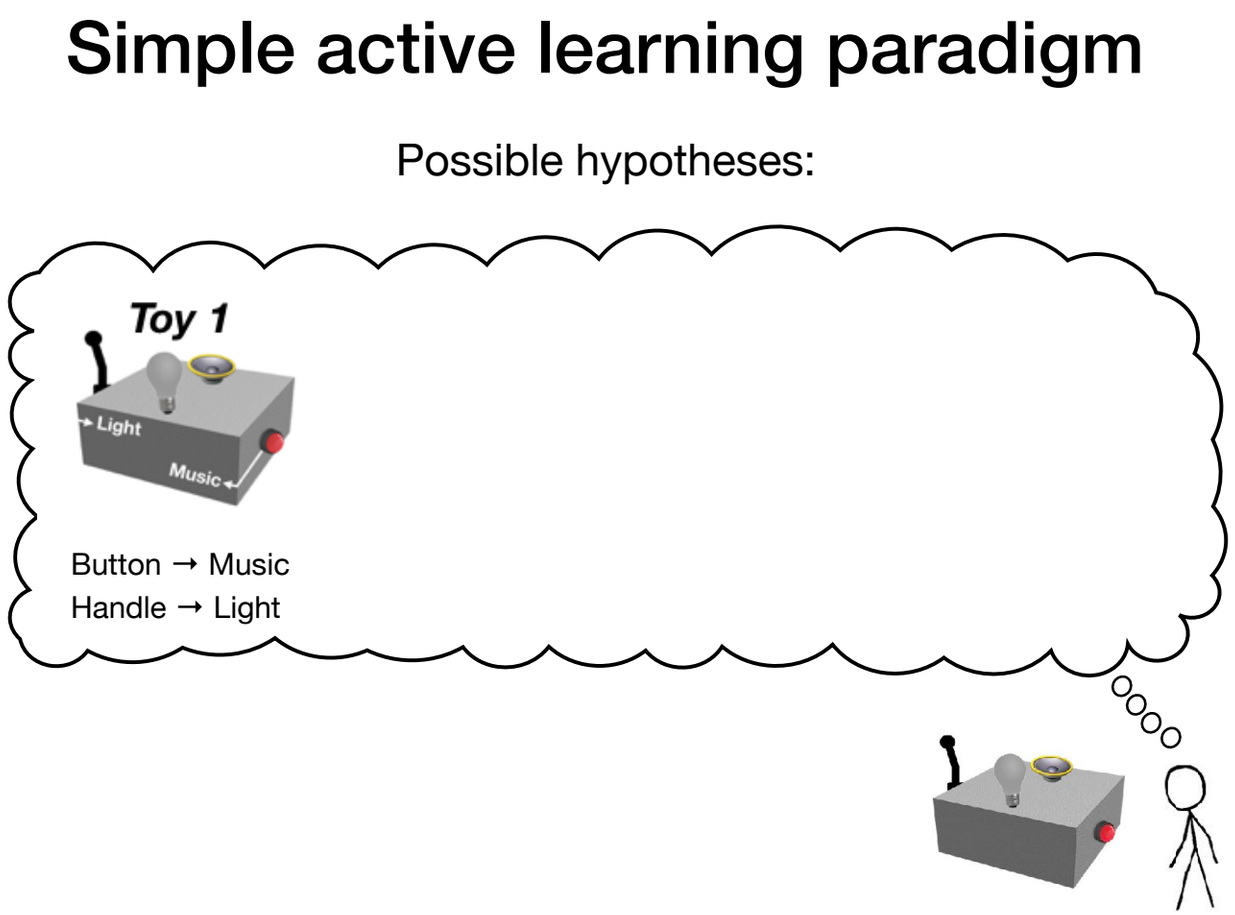
Simple active learning paradigm

Possible hypotheses:

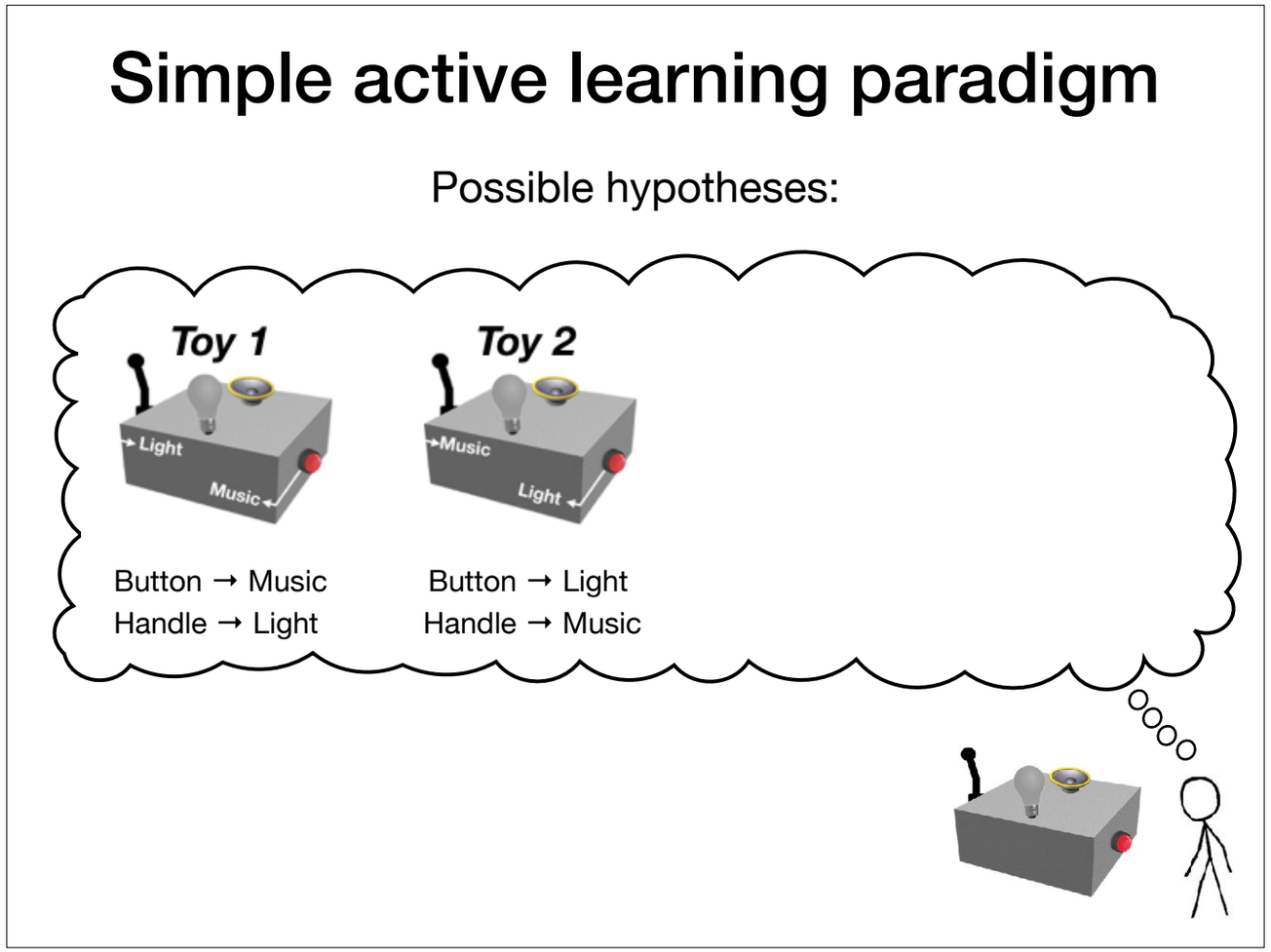


But imagine that you know that this toy is one of the three possible toys in this world. For **Toy** 1, pressing the button turns on the music, whereas pulling the handle turns on the light (and doing both will turn on both of those things.) For **Toy** 2, pressing the button turns on the light, whereas pulling the handle turns on the music. For **Toy** 3, Doing both actions at the same time will turn on both the music and light, but either of the single actions on its own will produce no effect.

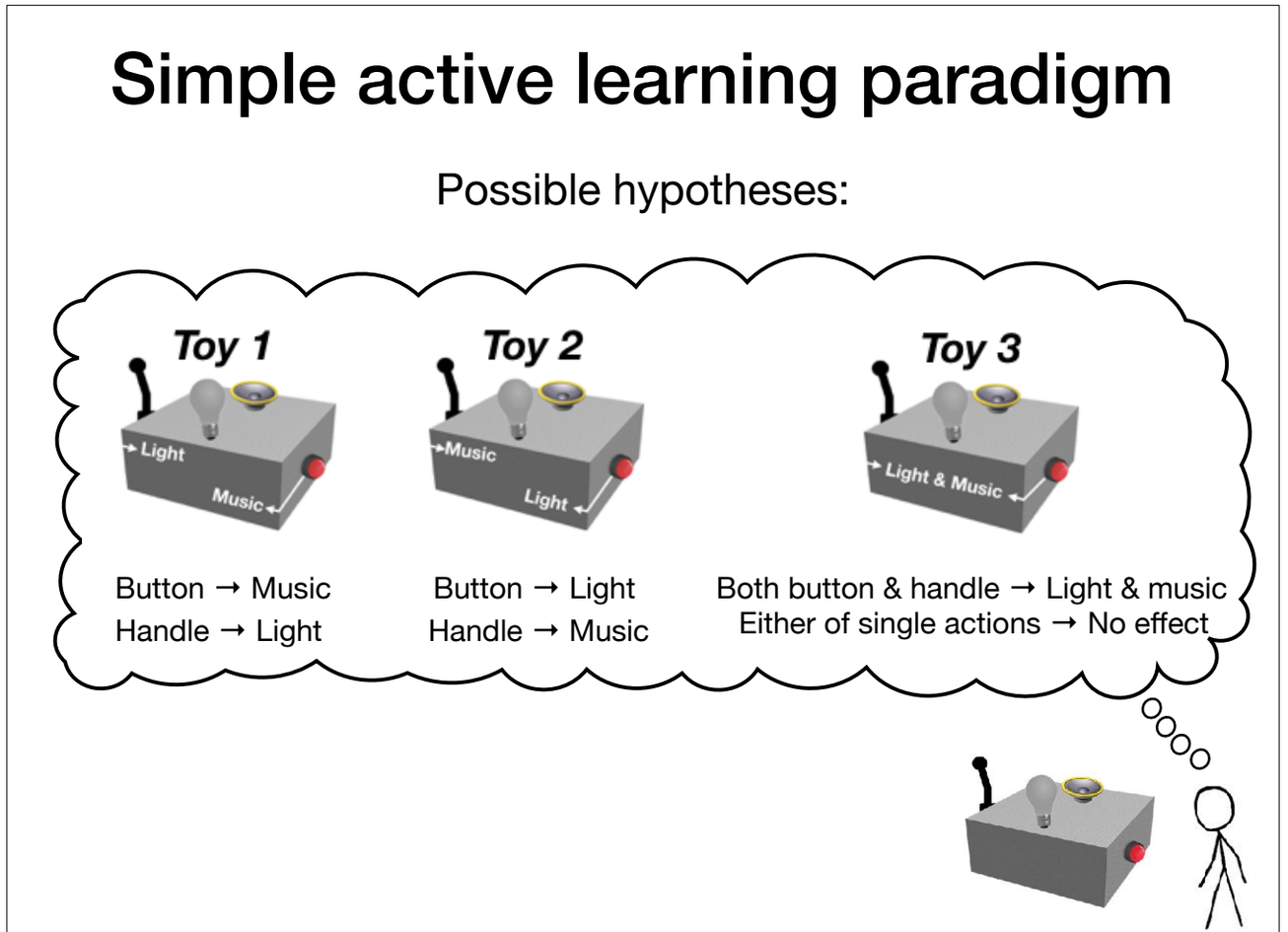
Simple active learning paradigm



But imagine that you know that this toy is one of the three possible toys in this world. For **Toy** 1, pressing the button turns on the music, whereas pulling the handle turns on the light (and doing both will turn on both of those things.) For **Toy** 2, pressing the button turns on the light, whereas pulling the handle turns on the music. For **Toy** 3, Doing both actions at the same time will turn on both the music and light, but either of the single actions on its own will produce no effect.

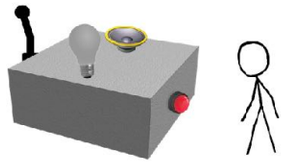


But imagine that you know that this toy is one of the three possible toys in this world. For **Toy** 1, pressing the button turns on the music, whereas pulling the handle turns on the light (and doing both will turn on both of those things.) For **Toy** 2, pressing the button turns on the light, whereas pulling the handle turns on the music. For **Toy** 3, Doing both actions at the same time will turn on both the music and light, but either of the single actions on its own will produce no effect.



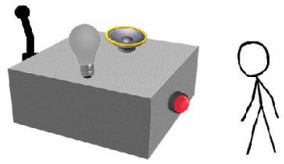
But imagine that you know that this toy is one of the three possible toys in this world. For **Toy** 1, pressing the button turns on the music, whereas pulling the handle turns on the light (and doing both will turn on both of those things.) For **Toy** 2, pressing the button turns on the light, whereas pulling the handle turns on the music. For **Toy** 3, Doing both actions at the same time will turn on both the music and light, but either of the single actions on its own will produce no effect.

Simple active learning paradigm



Then, given these hypotheses, it leads to an implication about the possible actions you can take depending on your goal. If there was only **ONE** chance to do an intervention on the toy, and you could either take one of the **single** actions (so to pull on the handle only or to press the button only), or take **both** actions at the same time, then you're faced with a dilemma because these actions have different pros and cons.

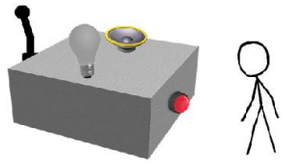
Simple active learning paradigm



If there is only **one** chance for an intervention, should you choose to take:

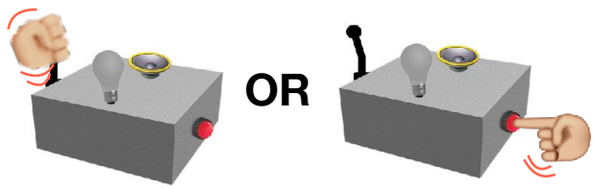
Then, given these hypotheses, it leads to an implication about the possible actions you can take depending on your goal. If there was only **ONE** chance to do an intervention on the toy, and you could either take one of the **single** actions (so to pull on the handle only or to press the button only), or take **both** actions at the same time, then you're faced with a dilemma because these actions have different pros and cons.

Simple active learning paradigm



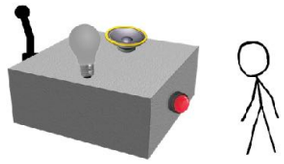
If there is only **one** chance for an intervention, should you choose to take:

a **single** action

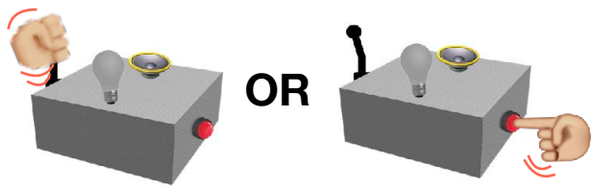
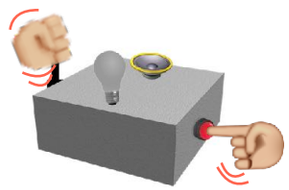


Then, given these hypotheses, it leads to an implication about the possible actions you can take depending on your goal. If there was only **ONE** chance to do an intervention on the toy, and you could either take one of the **single** actions (so to pull on the handle only or to press the button only), or take **both** actions at the same time, then you're faced with a dilemma because these actions have different pros and cons.

Simple active learning paradigm

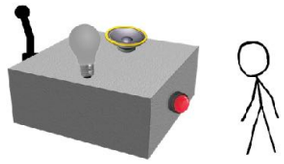


If there is only **one** chance for an intervention, should you choose to take:

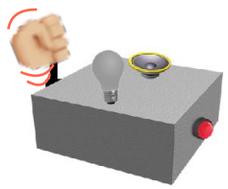
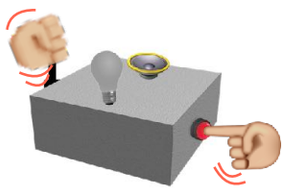
a **single** action	**both** actions
	

Then, given these hypotheses, it leads to an implication about the possible actions you can take depending on your goal. If there was only **ONE** chance to do an intervention on the toy, and you could either take one of the **single** actions (so to pull on the handle only or to press the button only), or take **both** actions at the same time, then you're faced with a dilemma because these actions have different pros and cons.

Simple active learning paradigm

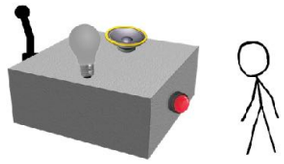


If there is only **one** chance for an intervention, should the learner choose to take:

a **single** action	**both** actions
	

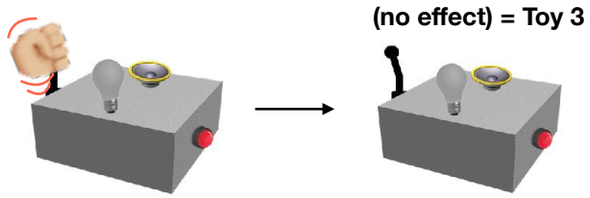
Taking a single action like pulling the handle only will **lead** to a full disambiguation, because you will definitely learn which of the three possible toys this is (for example, if you get no effect, then it's Toy 3 where you have to do both actions to produce the effects), but you risk having no immediate reward, just like in this case where there was no effect produced. On the other hand, if you take **both** actions, then you will definitely get an immediate rewarding effect, but this will not disambiguate which toy it is at all (because taking both actions on any of the three toys produces exactly the same outcome).

Simple active learning paradigm



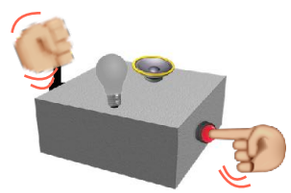
If there is only **one** chance for an intervention, should the learner choose to take:

a **single** action



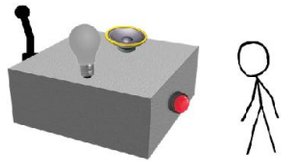
disambiguates (learning ↑), but risks no reward (activation ↓)

both actions



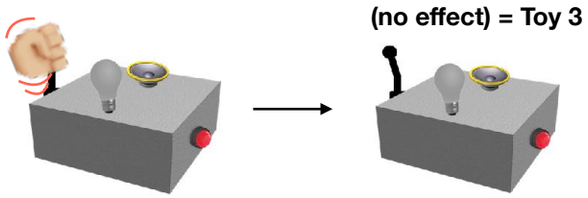
Taking a single action like pulling the handle only will **lead** to a full disambiguation, because you will definitely learn which of the three possible toys this is (for example, if you get no effect, then it's Toy 3 where you have to do both actions to produce the effects), but you risk having no immediate reward, just like in this case where there was no effect produced. On the other hand, if you take **both** actions, then you will definitely get an immediate rewarding effect, but this will not disambiguate which toy it is at all (because taking both actions on any of the three toys produces exactly the same outcome).

Simple active learning paradigm



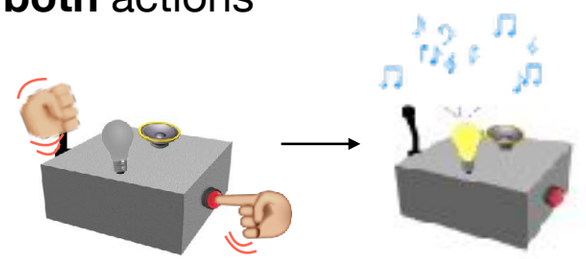
If there is only **one** chance for an intervention, should the learner choose to take:

a **single** action



disambiguates (learning ↑), but risks no reward (activation ↓)

both actions

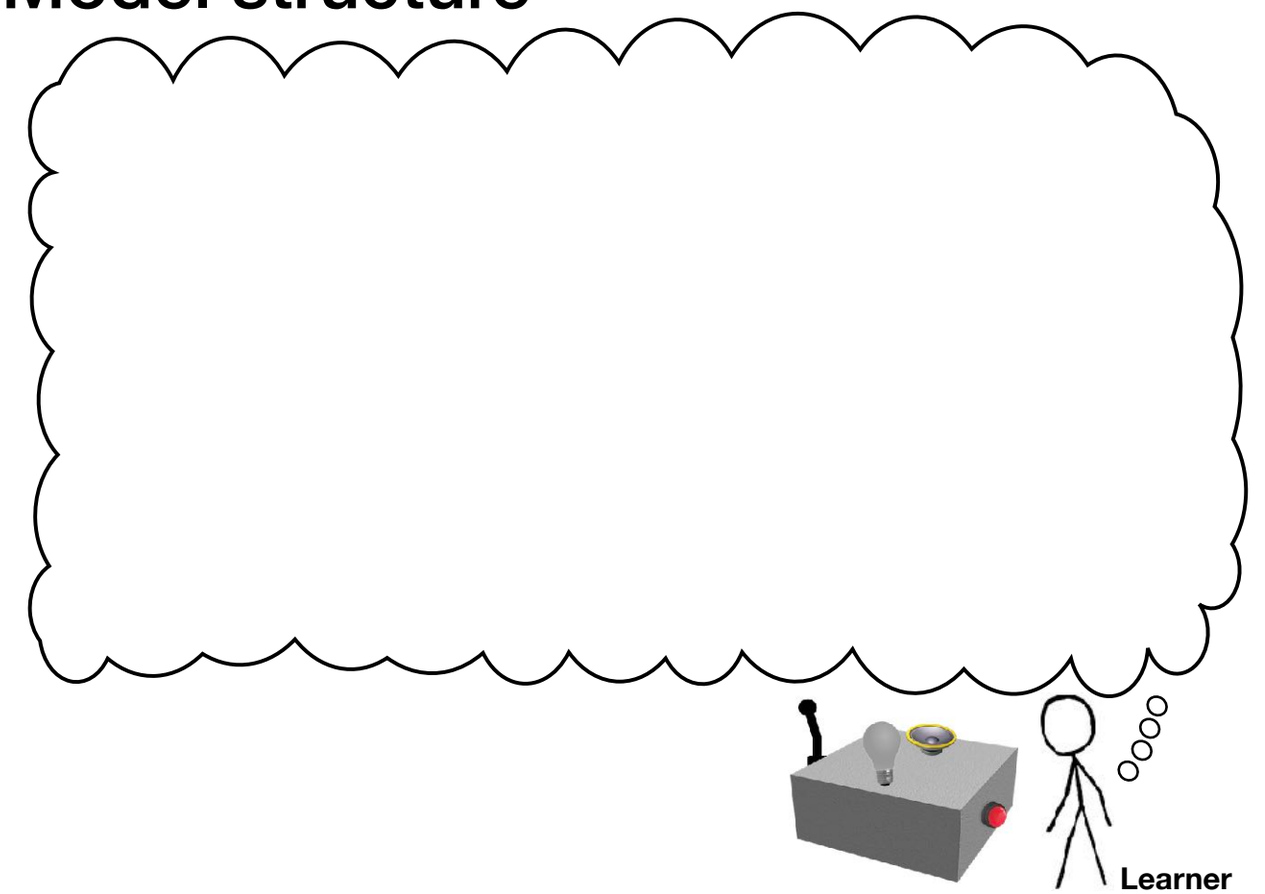


ensures immediate reward (activation ↑), but does not disambiguate (learning ↓)

Taking a single action like pulling the handle only will **lead** to a full disambiguation, because you will definitely learn which of the three possible toys this is (for example, if you get no effect, then it's Toy 3 where you have to do both actions to produce the effects), but you risk having no immediate reward, just like in this case where there was no effect produced. On the other hand, if you take **both** actions, then you will definitely get an immediate rewarding effect, but this will not disambiguate which toy it is at all (because taking both actions on any of the three toys produces exactly the same outcome).

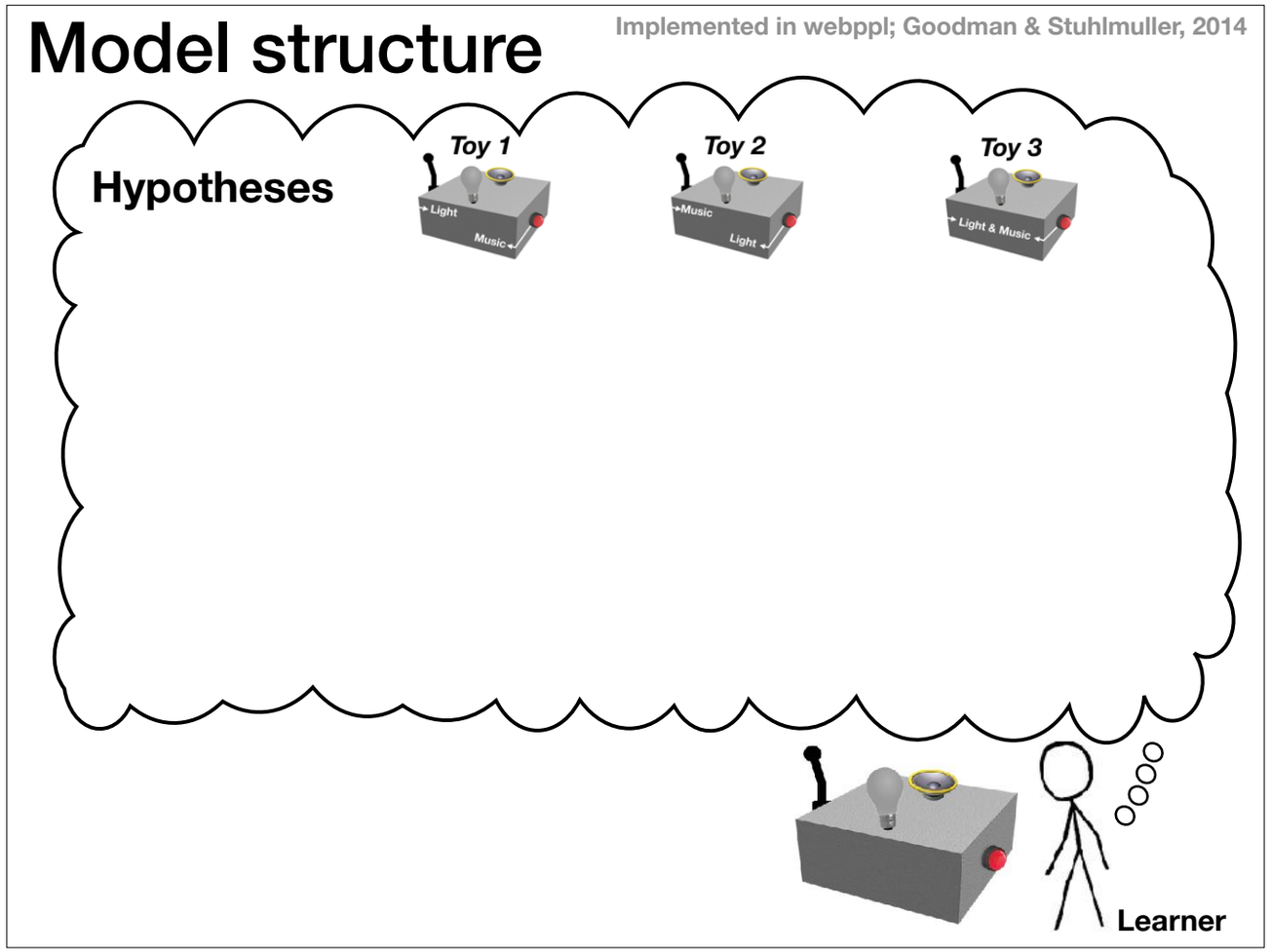
Model structure

Implemented in webppl; Goodman & Stuhlmuller, 2014



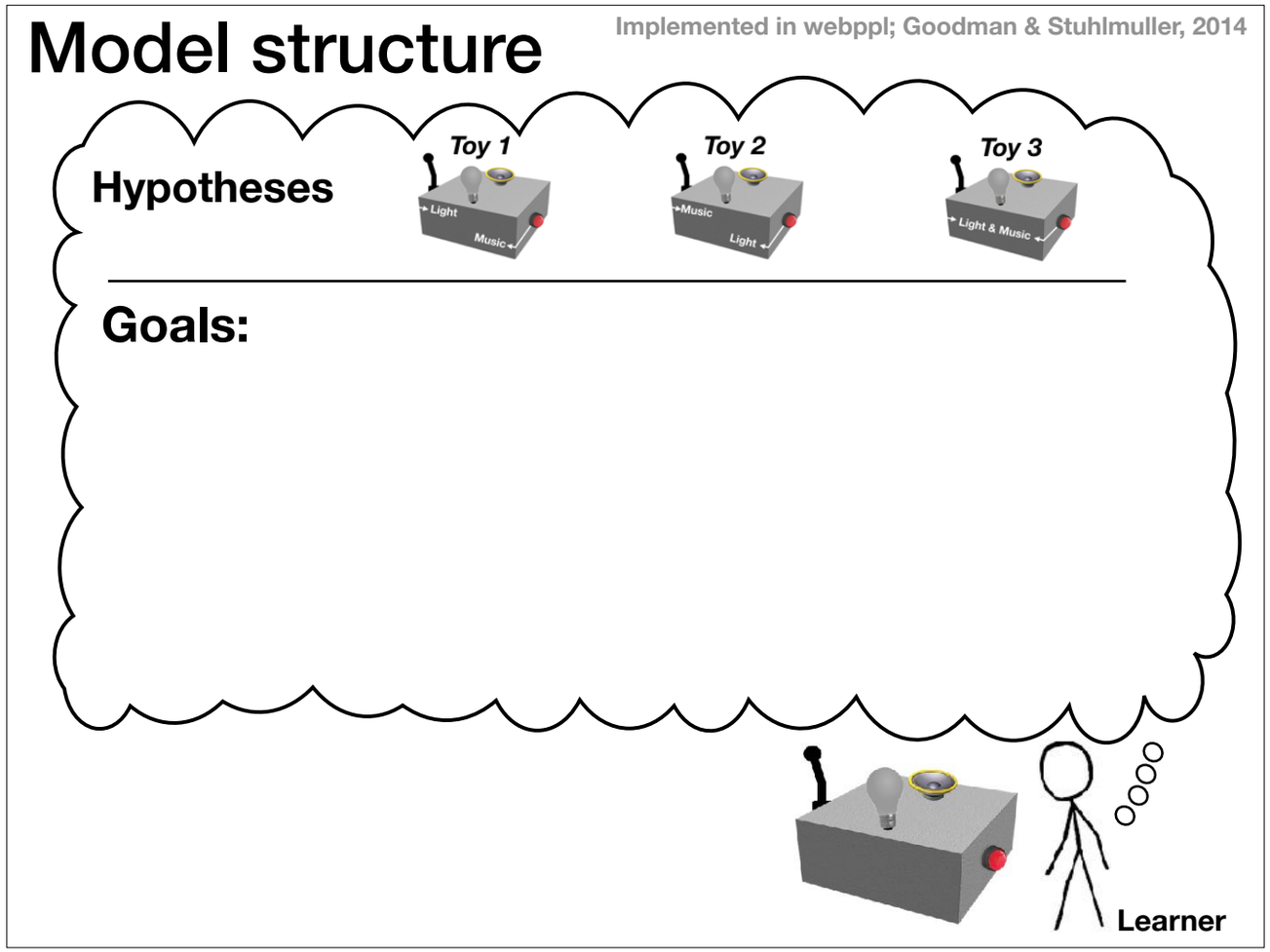
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



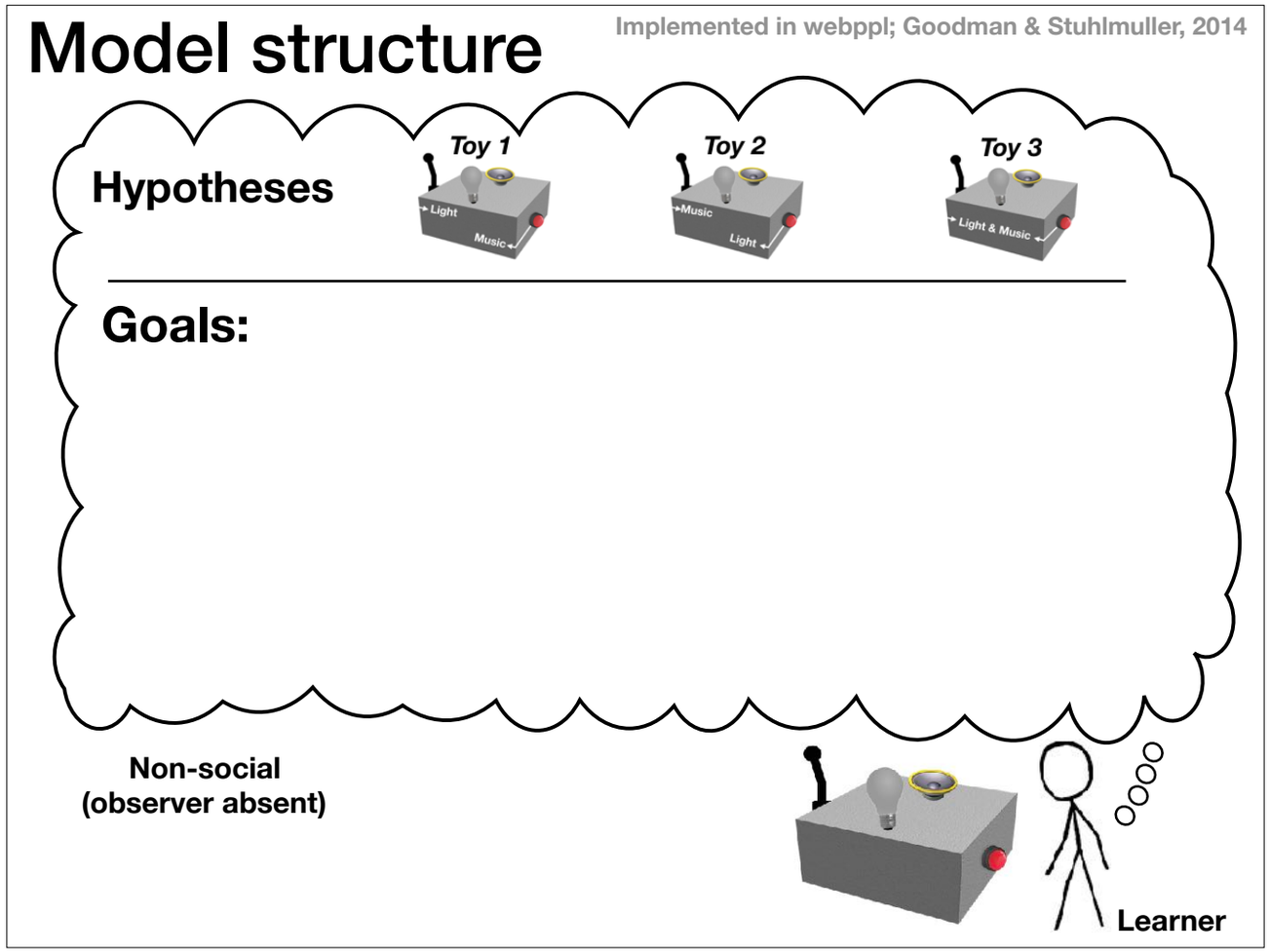
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



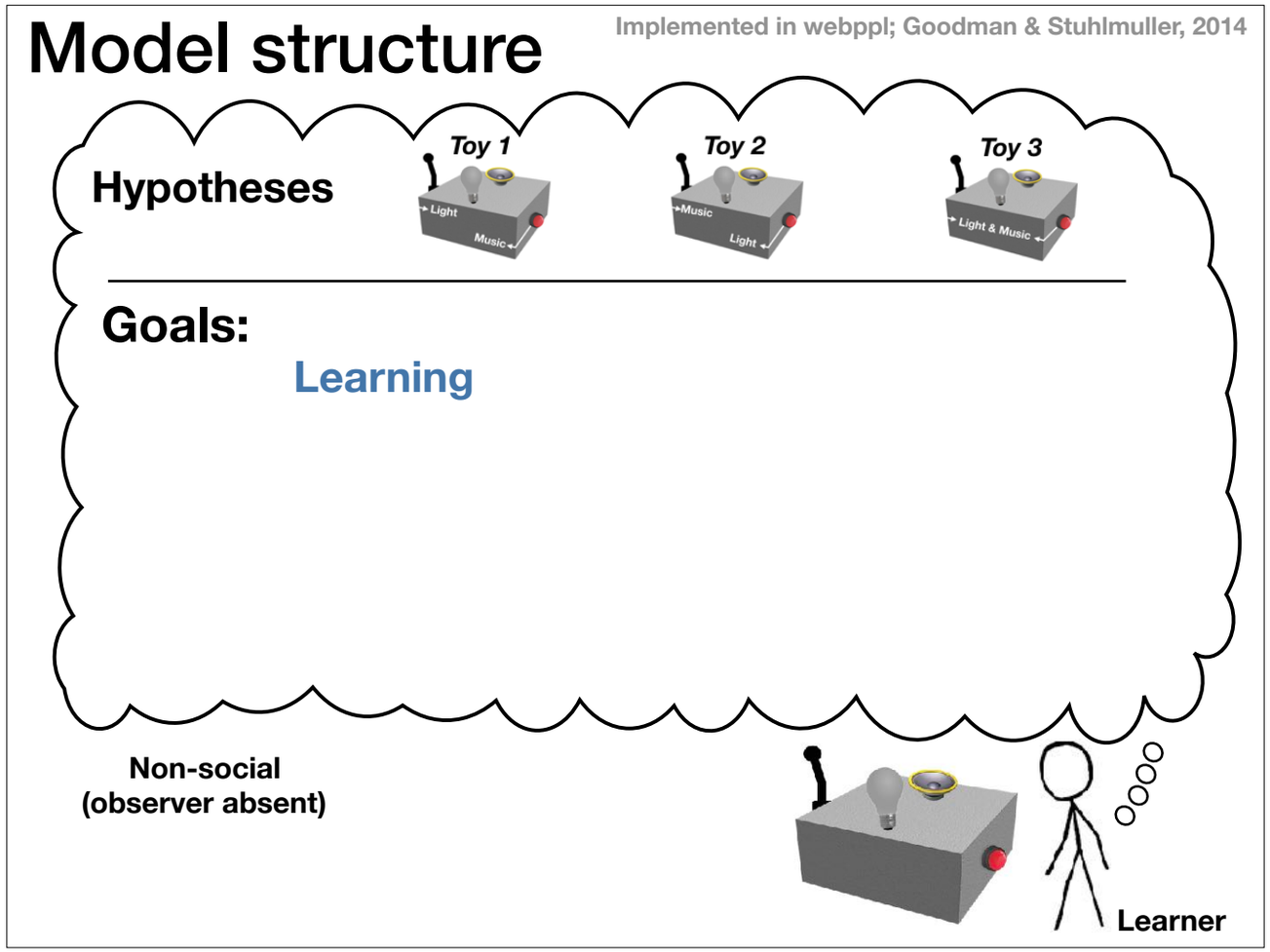
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



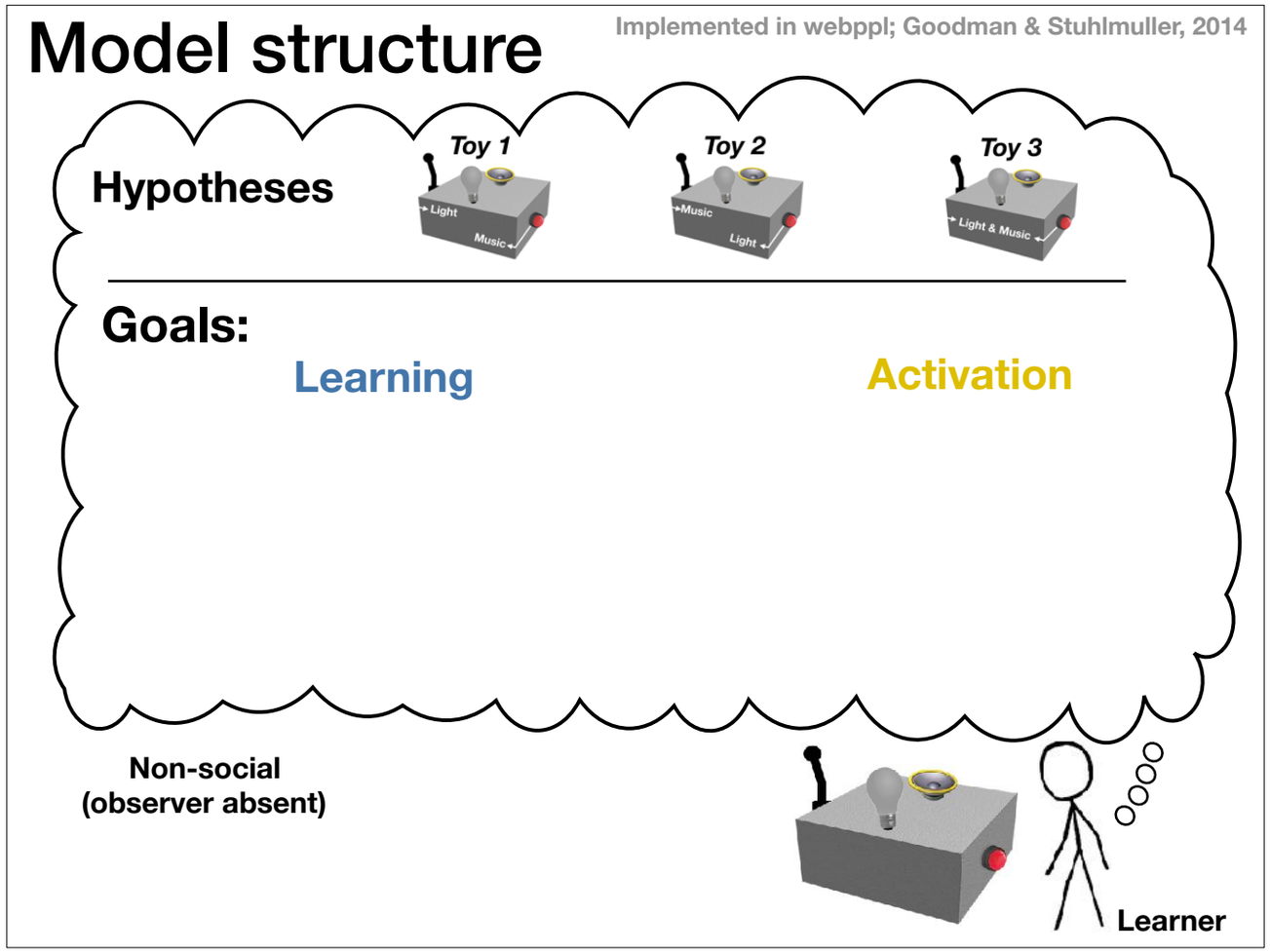
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



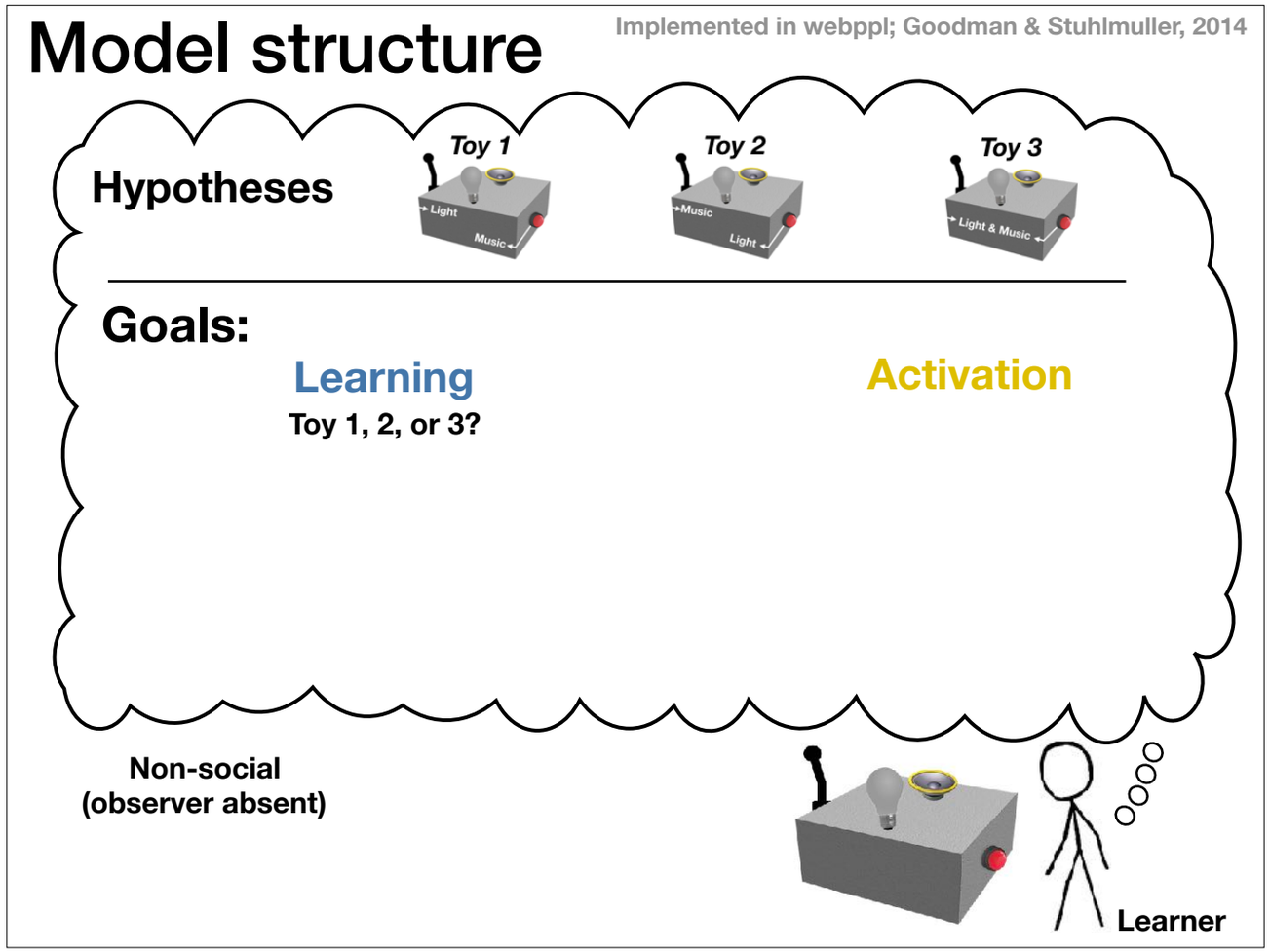
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



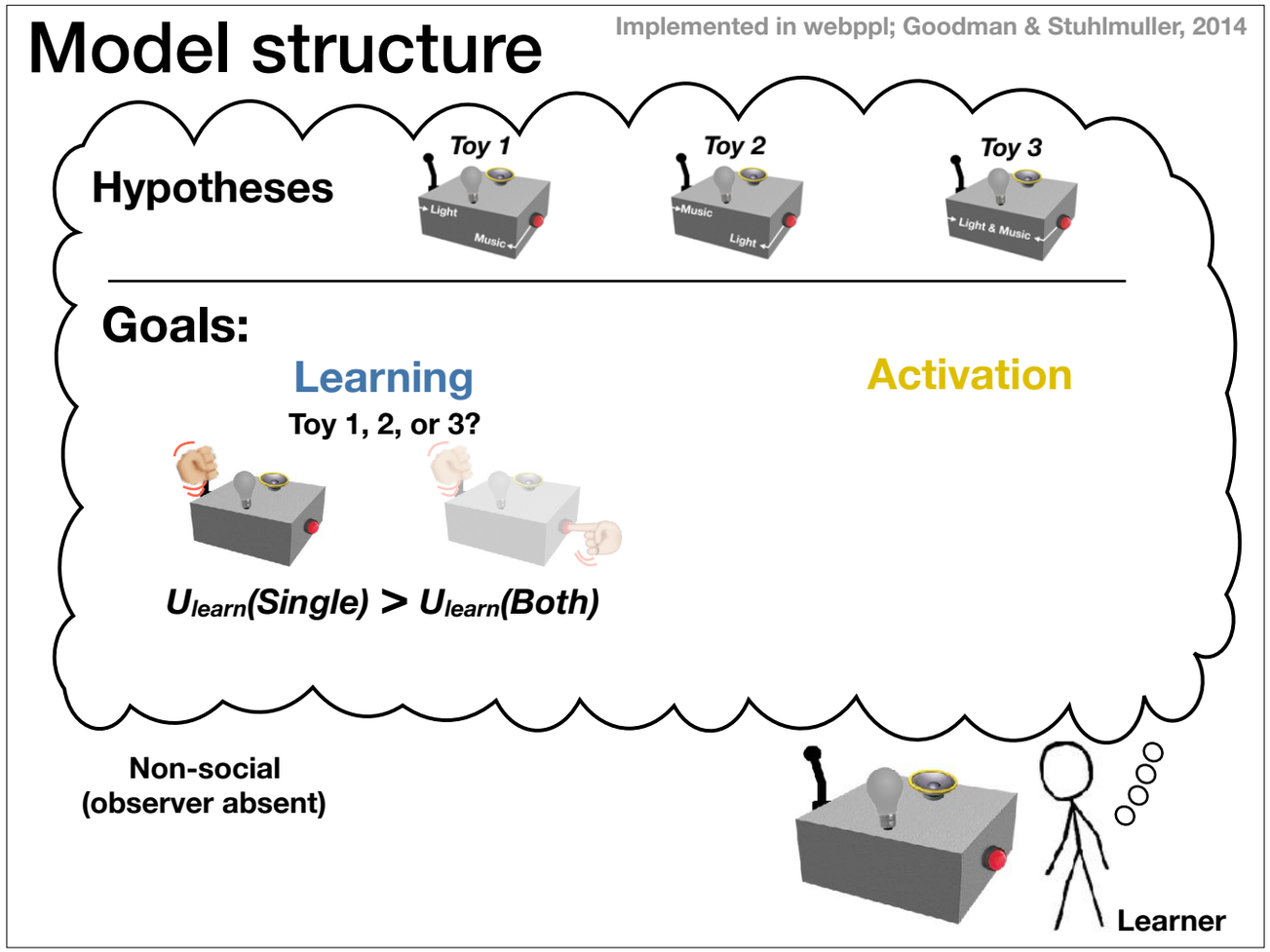
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



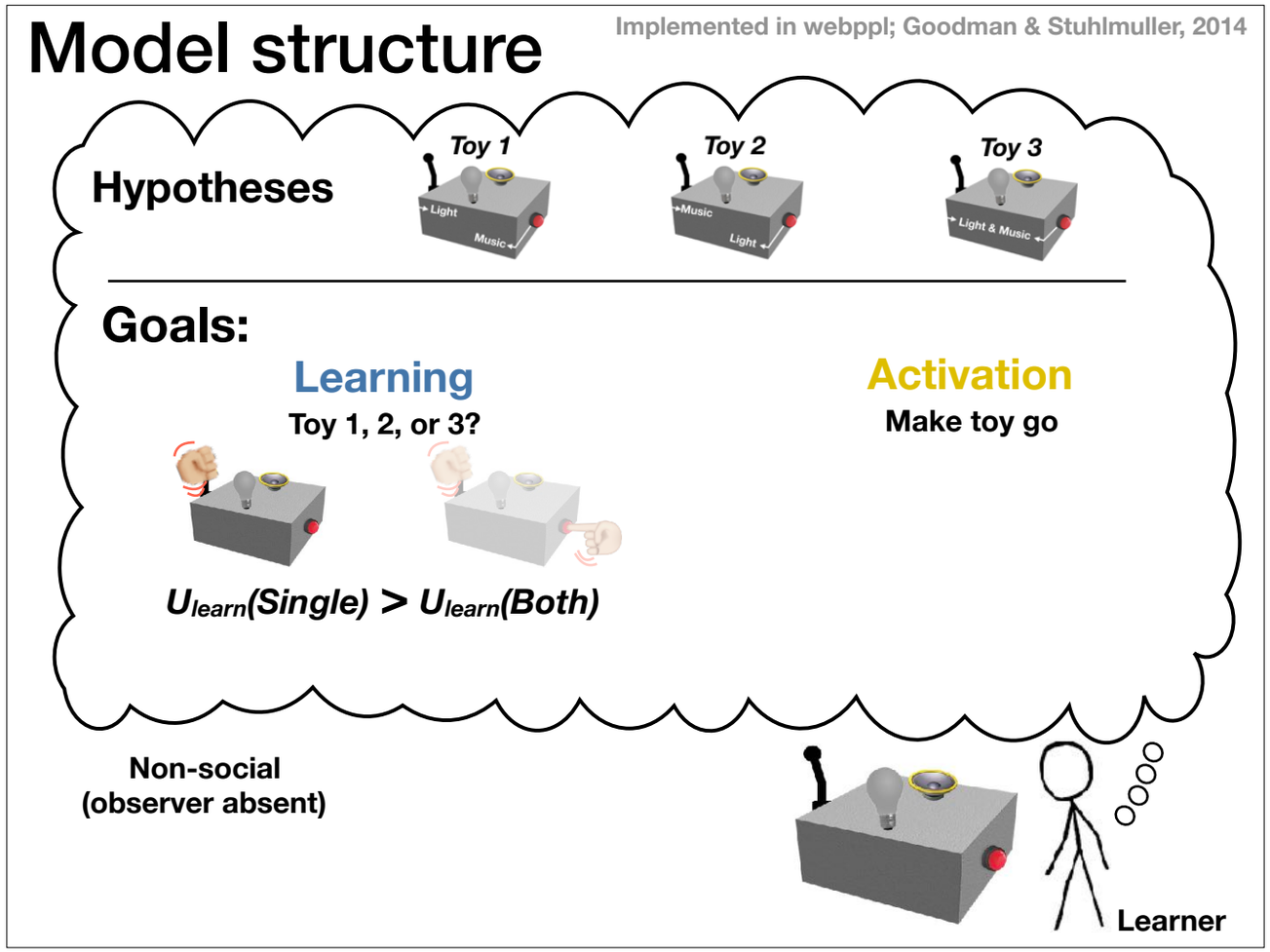
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



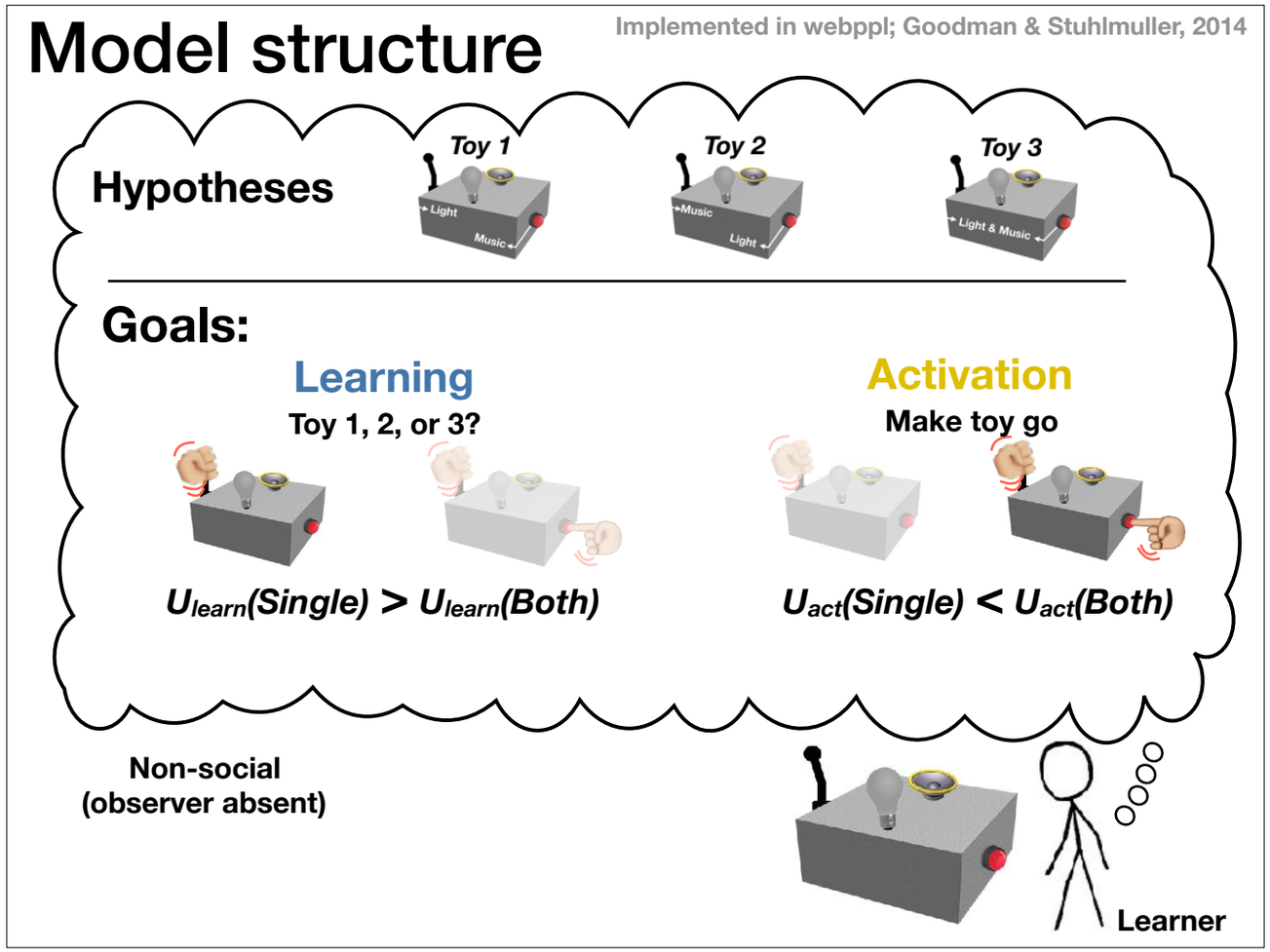
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



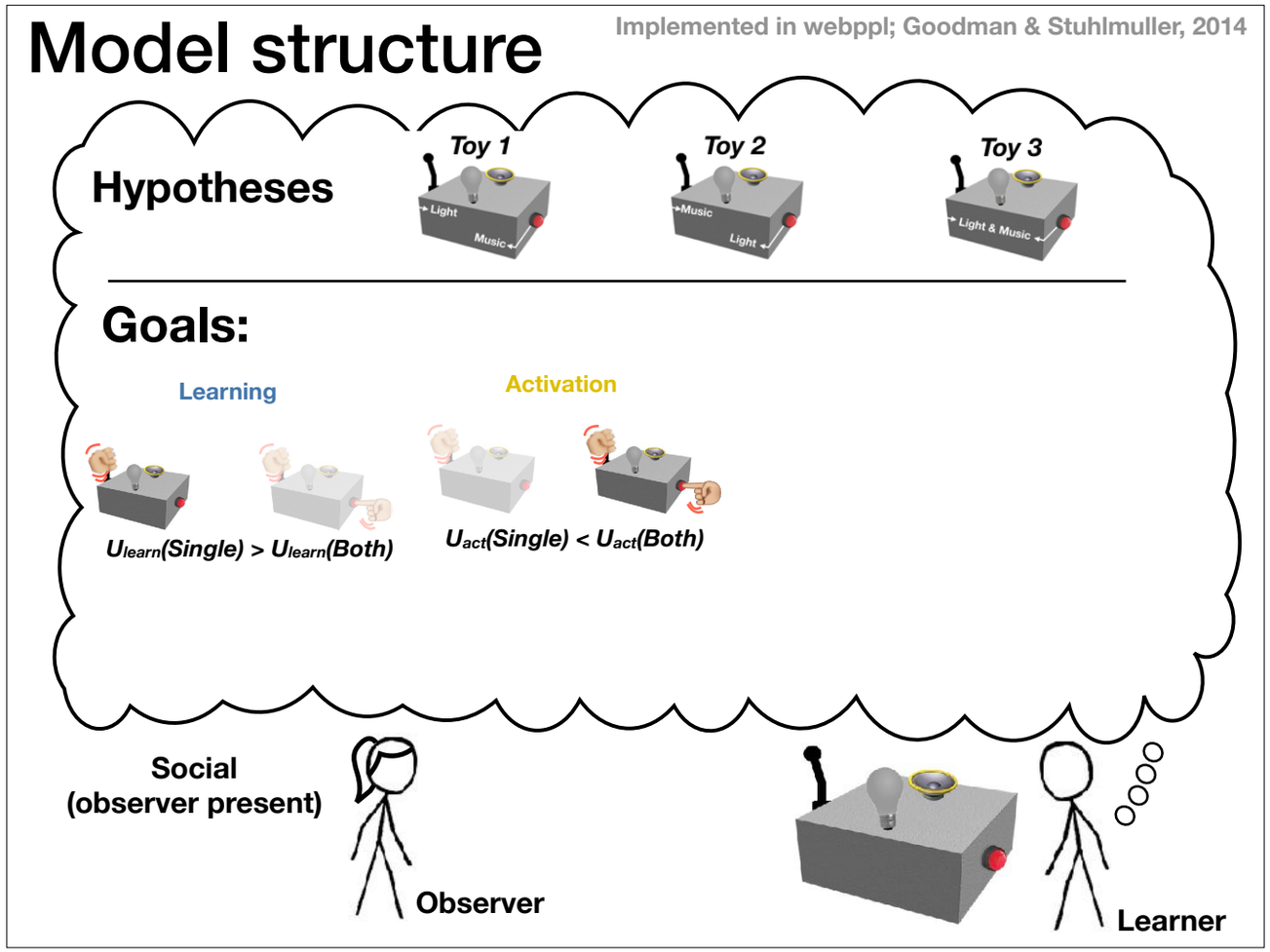
So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.

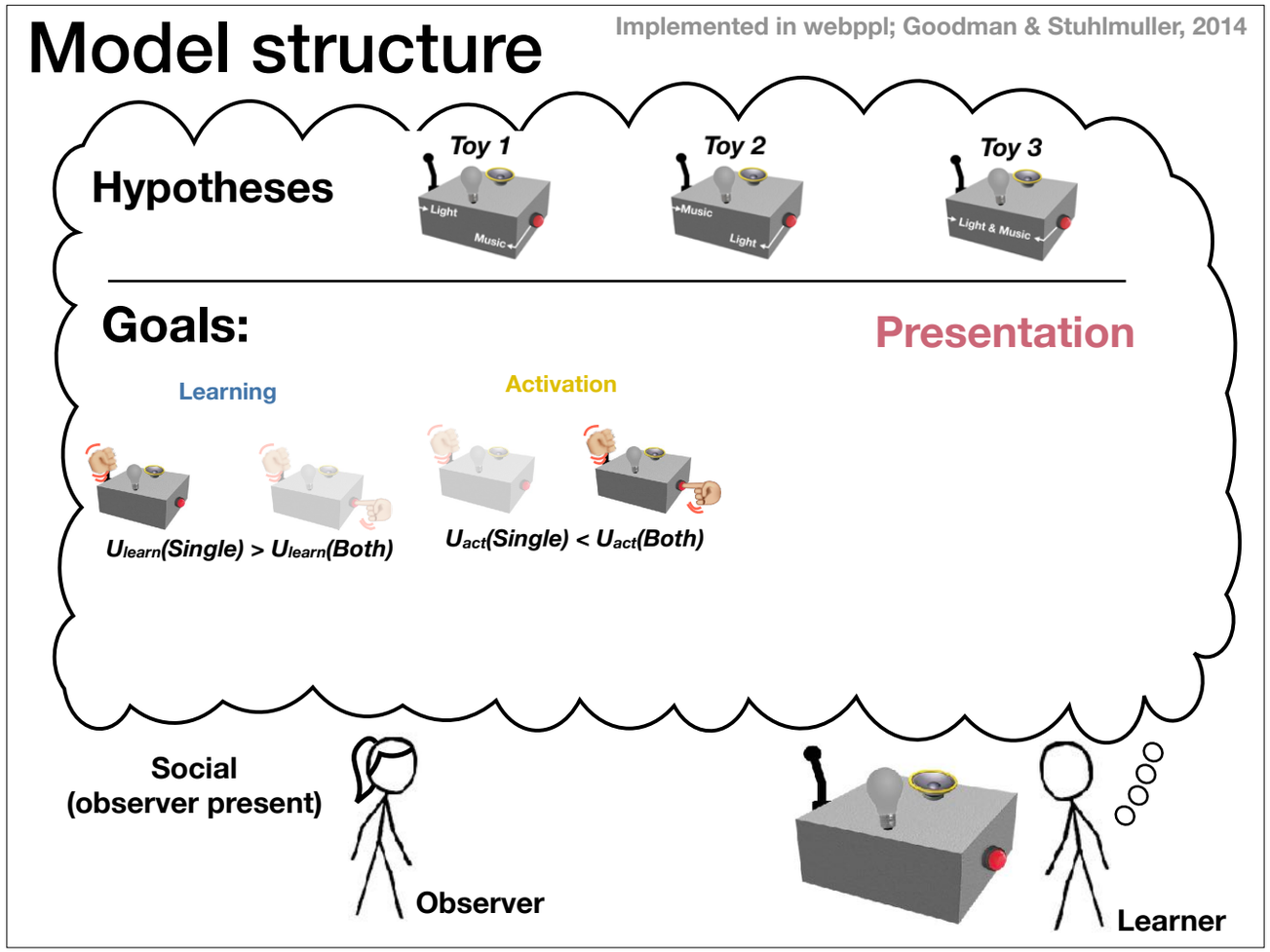


So how can we model what a learner might think in this kind of situation? First, we assume that the learner has **knowledge** about the three possible hypotheses. And the learner considers his **goals**. When there is **No** observer present, the learner considers two goals: a **learning** goal and **activation** goal. The learning goal has to do with **learning** which of the three hypotheses is true, and so maximizing information gain to reduce entropy, or the uncertainty about the toy mechanism. For this, the utility is higher for a **single** action than both actions because the single action is disambiguating.

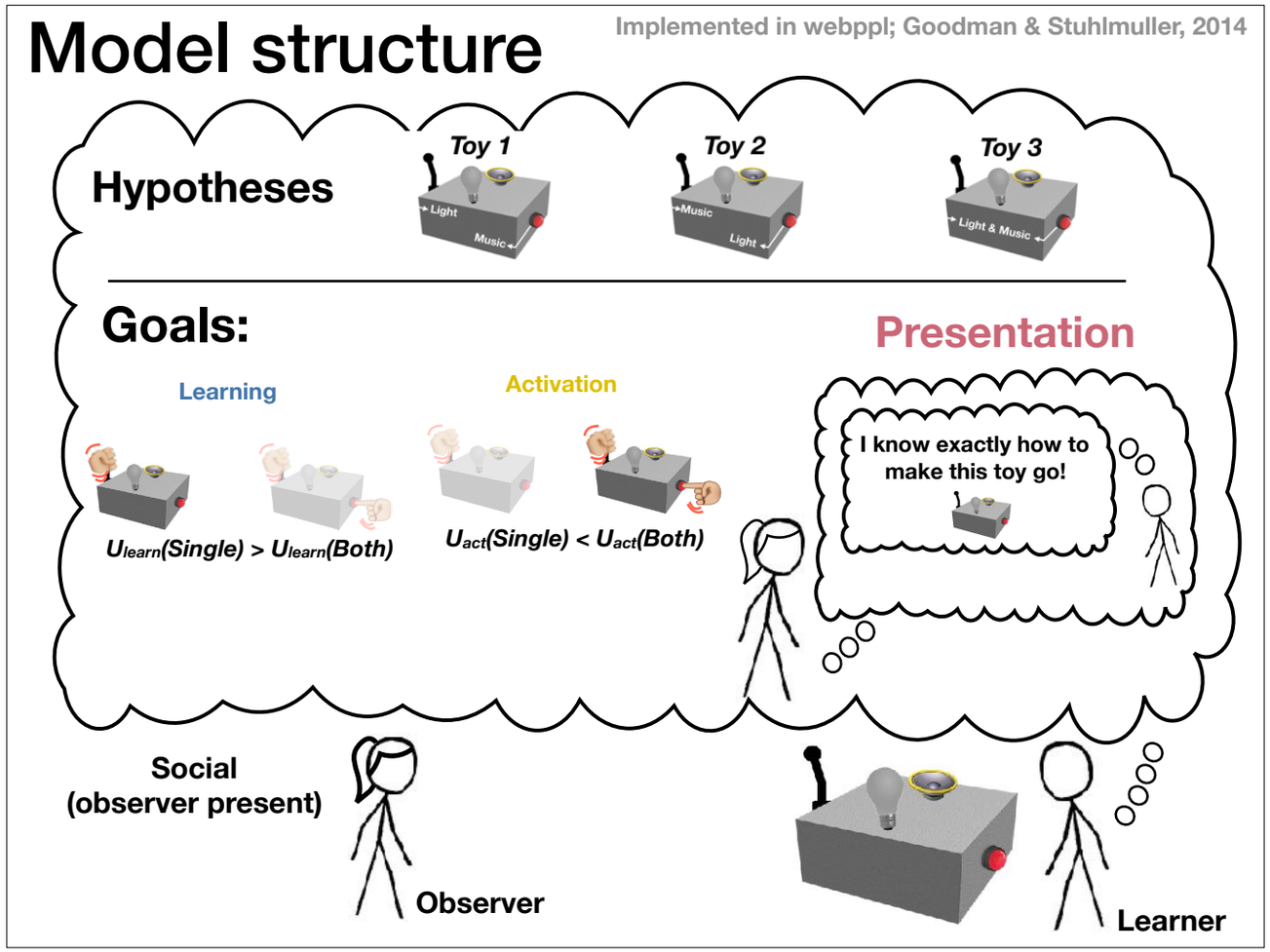
Activation utility is the utility of achieving an immediate reward outcome. In our paradigm, the learner gains utility from an immediate effect of music or light turning on, and so the expected activation utility is the likelihood of an effect given the action. So the utility is **higher** for both actions that will definitely produce an immediate effect.



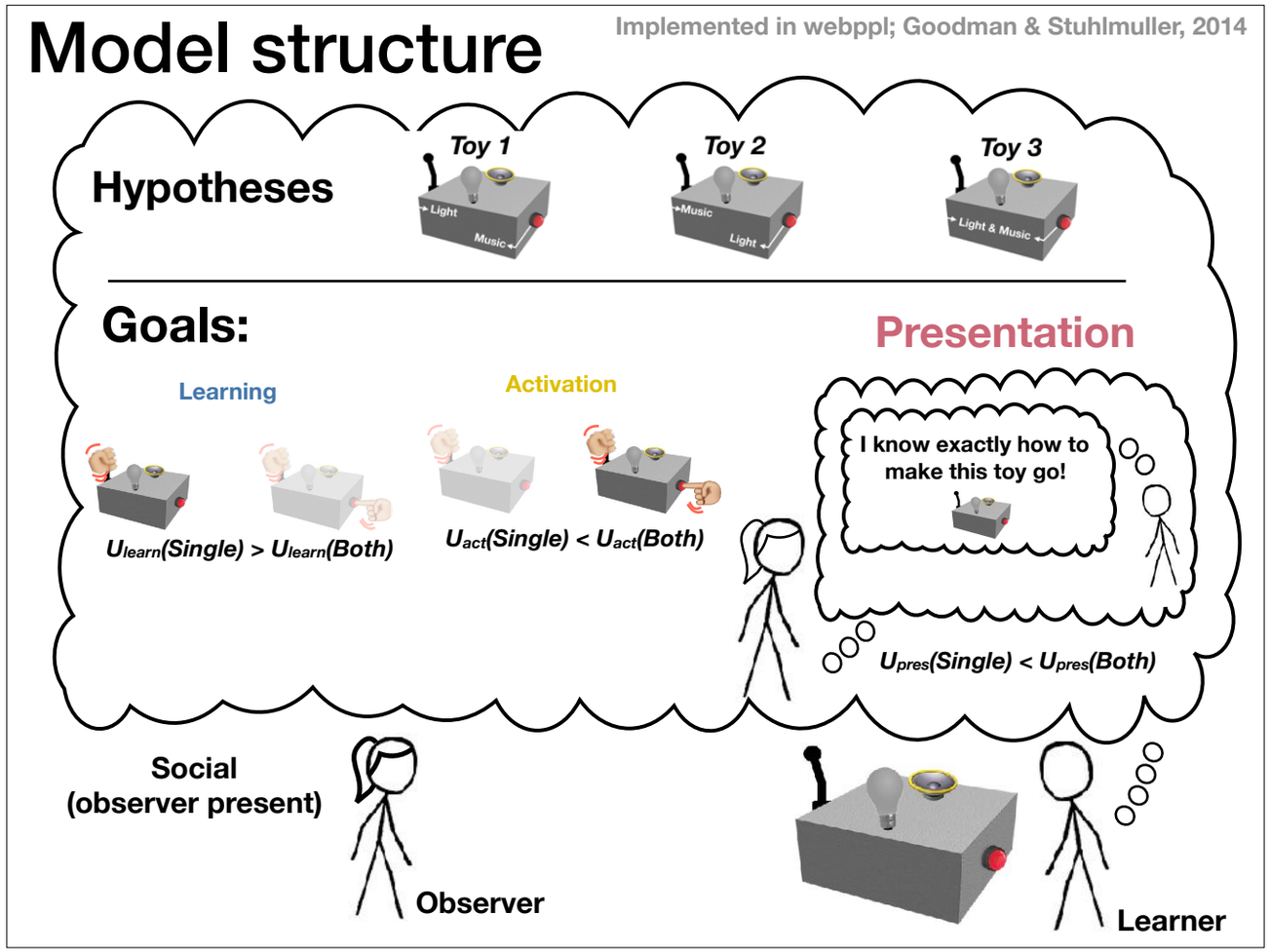
If an observer is present, the learner now considers a third utility, with has to do with his self-**presentation**. In our current model, this utility is recursive, and has to do with making the **observer** think that the learner is competent at making the toy go. Note that we're assuming that the observer is completely naive about the toy and possible hypotheses in this case.
Within our paradigm, because both actions ensure the maximum likelihood of producing an effect, the learner is predicted to **prefer** taking both actions.



If an observer is present, the learner now considers a third utility, with has to do with his self-**presentation**. In our current model, this utility is recursive, and has to do with making the **observer** think that the learner is competent at making the toy go. Note that we're assuming that the observer is completely naive about the toy and possible hypotheses in this case.
Within our paradigm, because both actions ensure the maximum likelihood of producing an effect, the learner is predicted to **prefer** taking both actions.



If an observer is present, the learner now considers a third utility, with has to do with his self-**presentation**. In our current model, this utility is recursive, and has to do with making the **observer** think that the learner is competent at making the toy go. Note that we're assuming that the observer is completely naive about the toy and possible hypotheses in this case.
Within our paradigm, because both actions ensure the maximum likelihood of producing an effect, the learner is predicted to **prefer** taking both actions.



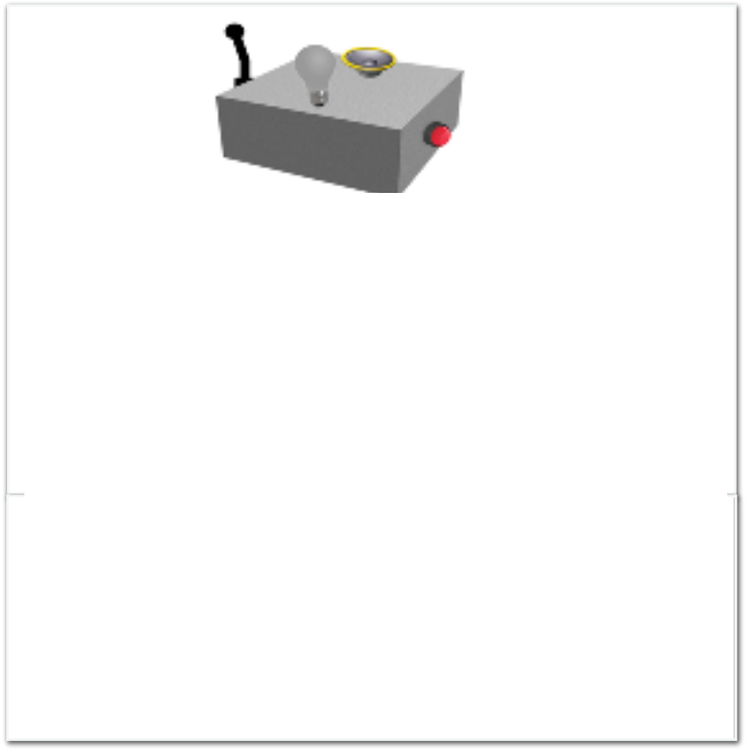
If an observer is present, the learner now considers a third utility, with has to do with his self-**presentation**. In our current model, this utility is recursive, and has to do with making the **observer** think that the learner is competent at making the toy go. Note that we're assuming that the observer is completely naive about the toy and possible hypotheses in this case.
Within our paradigm, because both actions ensure the maximum likelihood of producing an effect, the learner is predicted to **prefer** taking both actions.

Outline

- Active learning (+ social factors)
- Model of goal tradeoffs in active learning
  - Paradigm
  - Model structure
- **Experimental task**
  - Behavioral results
  - Model predictions

Next, we built an experimental task to test these predictions (which were pre-registered on osf).

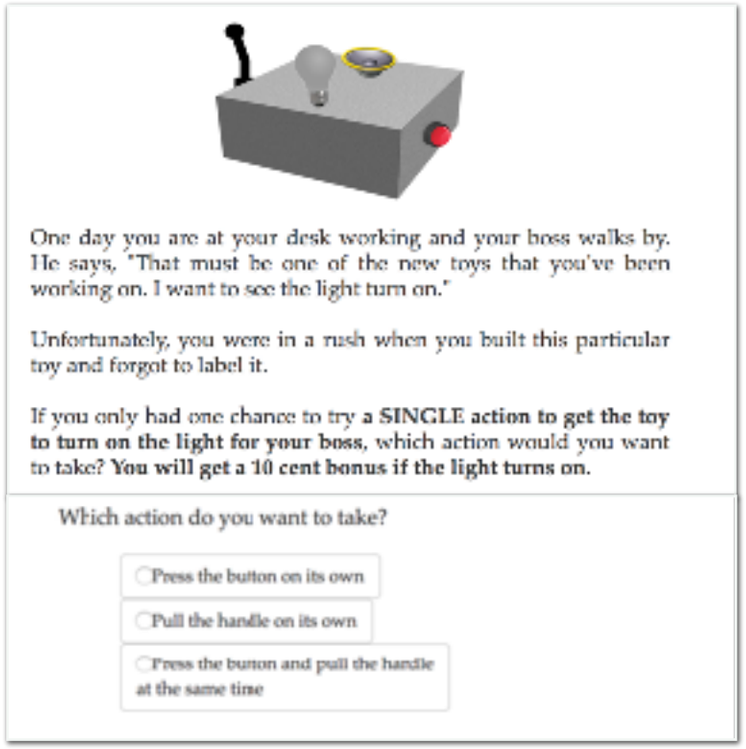
Method



MTurk, N = 325
(42-51 per condition)

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

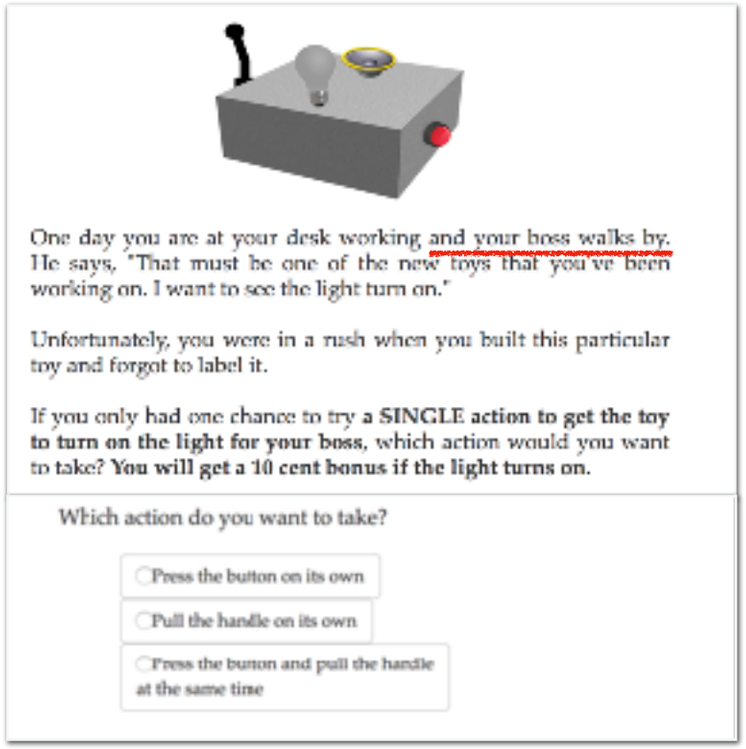
Method



MTurk, N = 325
(42-51 per condition)

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Method



MTurk, N = 325
(42-51 per condition)

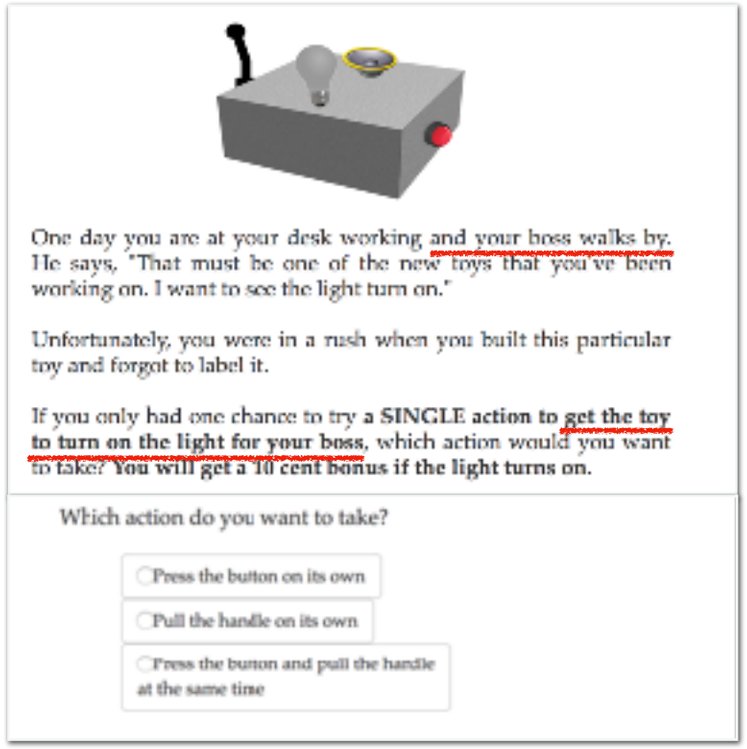
Manipulation:

Context:

- **social**: "your boss walks by"
- **non-social**: "you see one of your toys lying on the ground"

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Method



MTurk, N = 325
(42-51 per condition)

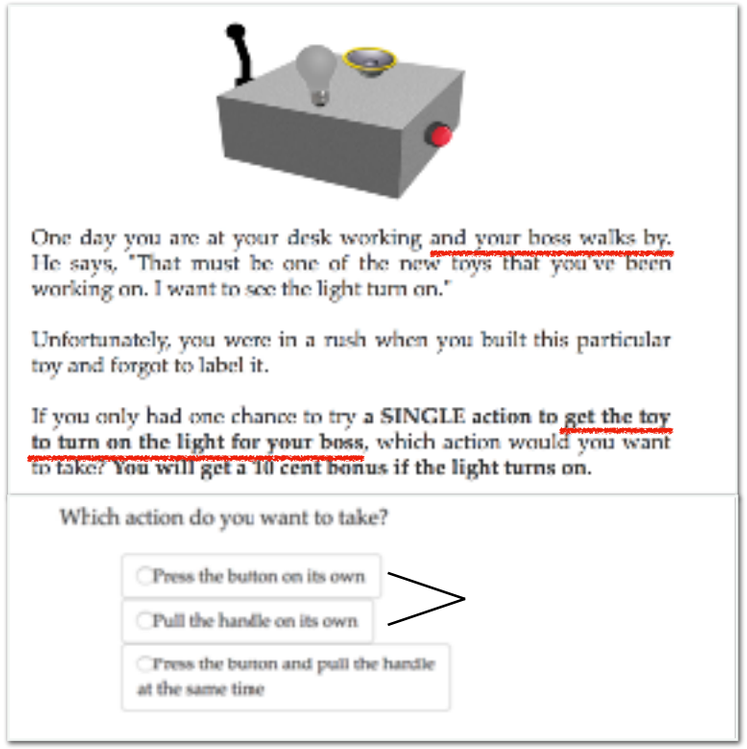
Manipulation:

Context:
- **social**: "your boss walks by"
- **non-social**: "you see one of your toys lying on the ground"

Goal:
- **learning**: "figure out the correct label for the toy"
- **activation**: "turn on the light"
- **presentation**: "impress your boss and show that you're competent"
- **no-goal**: "play with this toy"

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Method



MTurk, N = 325
(42-51 per condition)

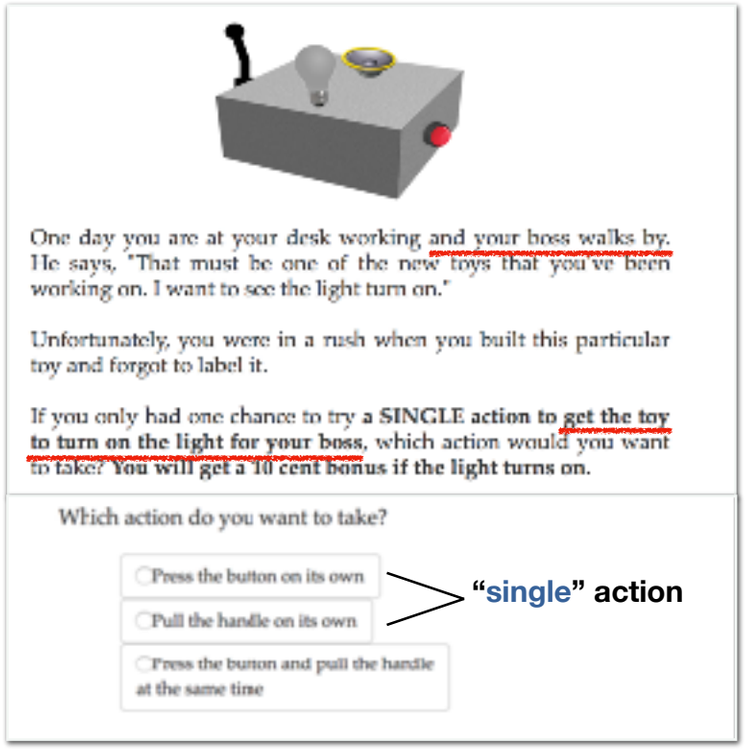
Manipulation:

Context:
- **social**: "your boss walks by"
- **non-social**: "you see one of your toys lying on the ground"

Goal:
- **learning**: "figure out the correct label for the toy"
- **activation**: "turn on the light"
- **presentation**: "impress your boss and show that you're competent"
- **no-goal**: "play with this toy"

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Method



MTurk, N = 325
(42-51 per condition)

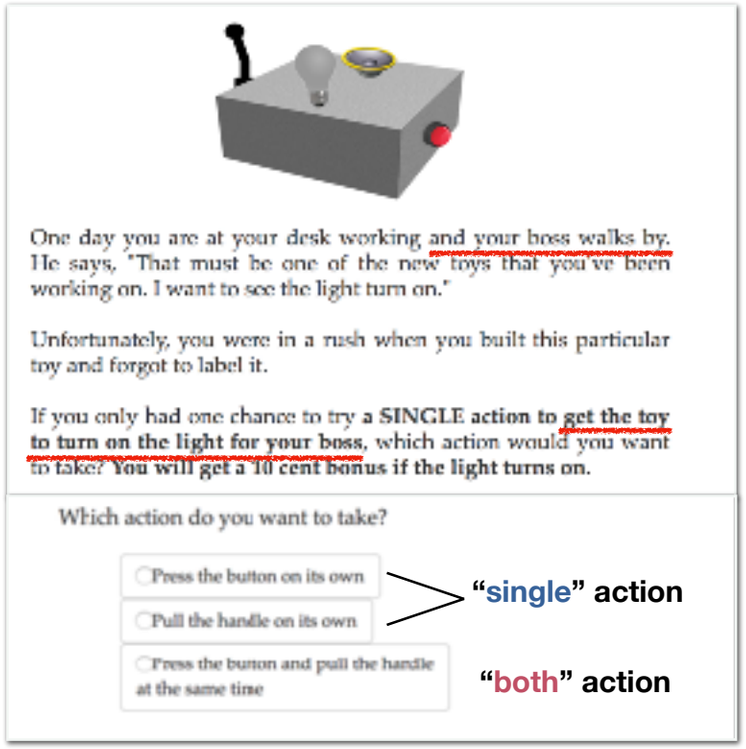
Manipulation:

Context:
- **social**: "your boss walks by"
- **non-social**: "you see one of your toys lying on the ground"

Goal:
- **learning**: "figure out the correct label for the toy"
- **activation**: "turn on the light"
- **presentation**: "impress your boss and show that you're competent"
- **no-goal**: "play with this toy"

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Method



MTurk, N = 325
(42-51 per condition)

Manipulation:

Context:
- **social**: "your boss walks by"
- **non-social**: "you see one of your toys lying on the ground"

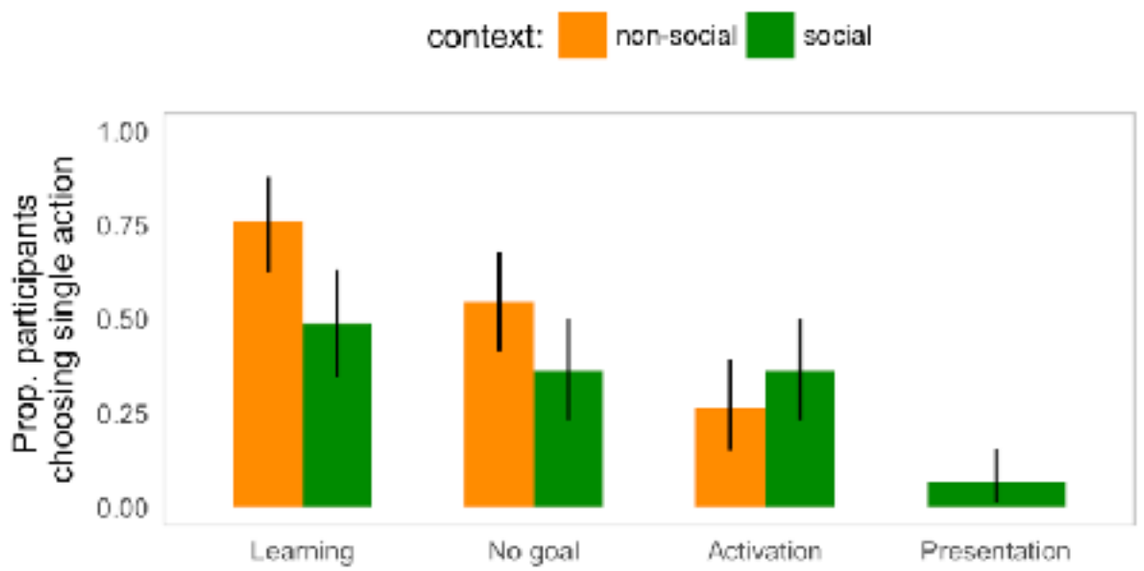
Goal:
- **learning**: "figure out the correct label for the toy"
- **activation**: "turn on the light"
- **presentation**: "impress your boss and show that you're competent"
- **no-goal**: "play with this toy"

Our task was conducted online on Mechanical Turk, and we asked participants to imagine that they were toy developers who designed a toy, and explained the three different possible toys to participants, and then presented a scenario such as **this**: One day… and asked them for their choice between the three options: …
We manipulated two things in this task: first, we varied the **context** such that it could be social with a boss present, or non-social with no presence of another person. second, we manipulated the **goal** description, such that the goal assigned to the participant could be: learning, or to figure out…, activation, or to turn on …, presentation, or to impress…, or no-goal, and to just play with the toy. And we used a 10 cent bonus to incentivize.
In terms of coding the action choices, because either of the single actions is fully disambiguating, we looked at two of **these** actions together as one category, versus "**both**" actions as the other category.

Results: Action choices

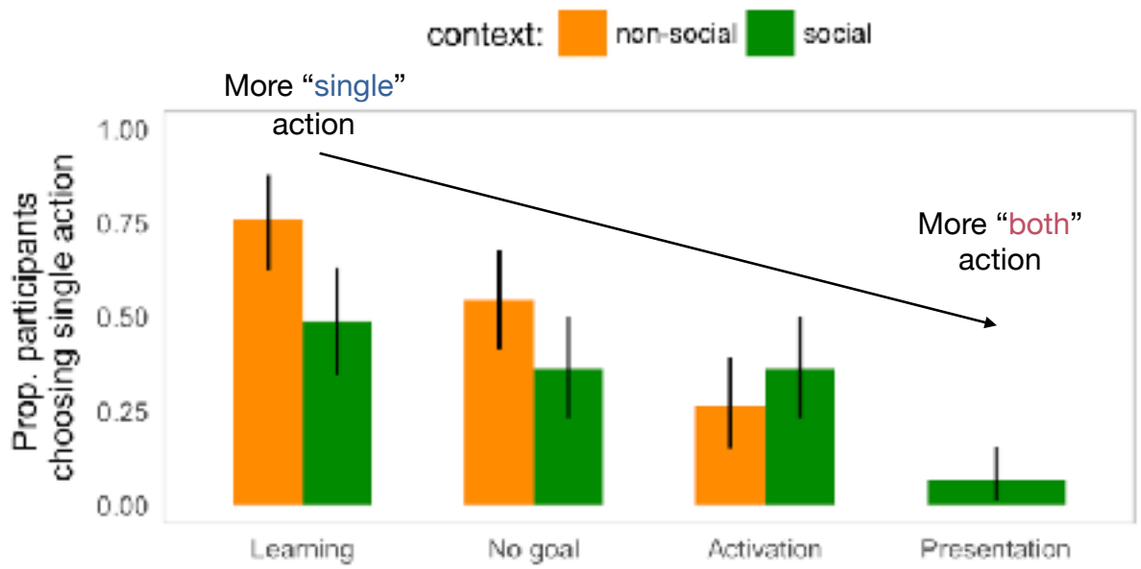
So here I'll show the results from participants' action choices. The X-axis shows the different goal conditions, and color represents social vs. non-social context. Y-axis shows the proportion of participants who chose the single action — so the action that is disambiguating but risking no immediate effect.
And here are the **results**. The first thing to notice is that participants **chose** more single actions in the order of learning condition, no-goal, activation, and presentation, which is consistent with what we had predicted. So having a learning goal leads participants to focus on more informative actions, whereas activation and presentation goals lead them to try to produce effects, and no-goal somewhere in between.
Second, there was an overall effect of **context**, such that social context yielded more immediately rewarding both actions than the non-social context, consistent with our idea that the presence of an observer would facilitate focus on immediate rewards more in this particular situation.
Third, there was an interaction between context and goal such that there was an effect of context on action choice within learning and no-goal conditions but not for activation goal.

Results: Action choices



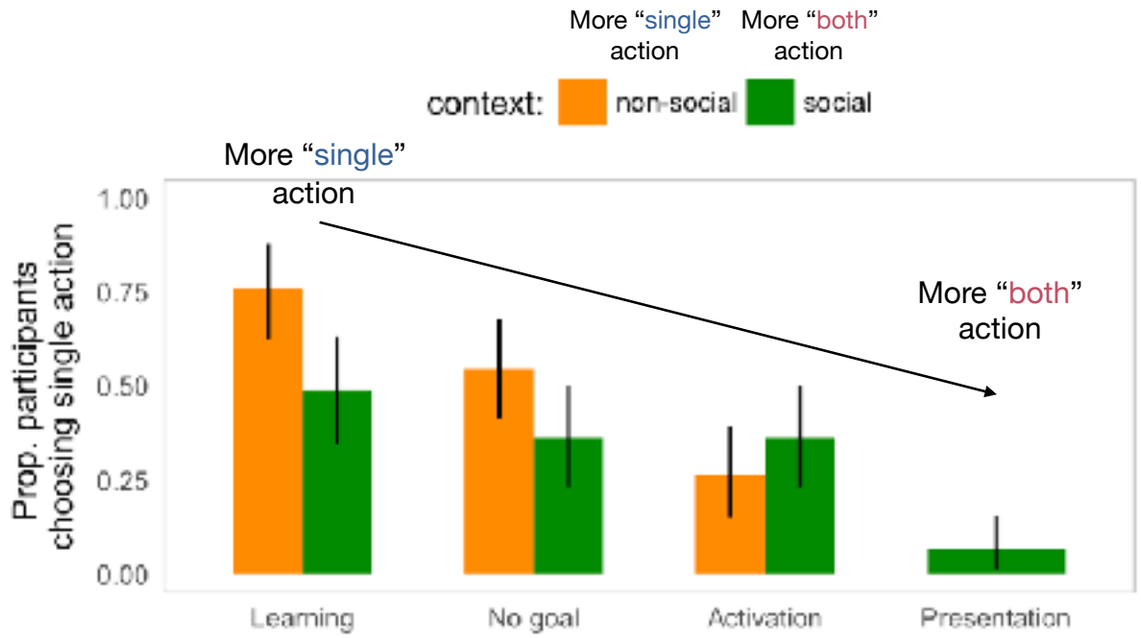
So here I'll show the results from participants' action choices. The X-axis shows the different goal conditions, and color represents social vs. non-social context. Y-axis shows the proportion of participants who chose the single action — so the action that is disambiguating but risking no immediate effect.
And here are the **results**. The first thing to notice is that participants **chose** more single actions in the order of learning condition, no-goal, activation, and presentation, which is consistent with what we had predicted. So having a learning goal leads participants to focus on more informative actions, whereas activation and presentation goals lead them to try to produce effects, and no-goal somewhere in between.
Second, there was an overall effect of **context**, such that social context yielded more immediately rewarding both actions than the non-social context, consistent with our idea that the presence of an observer would facilitate focus on immediate rewards more in this particular situation.
Third, there was an interaction between context and goal such that there was an effect of context on action choice within learning and no-goal conditions but not for activation goal.

Results: Action choices



So here I'll show the results from participants' action choices. The X-axis shows the different goal conditions, and color represents social vs. non-social context. Y-axis shows the proportion of participants who chose the single action — so the action that is disambiguating but risking no immediate effect.
And here are the **results**. The first thing to notice is that participants **chose** more single actions in the order of learning condition, no-goal, activation, and presentation, which is consistent with what we had predicted. So having a learning goal leads participants to focus on more informative actions, whereas activation and presentation goals lead them to try to produce effects, and no-goal somewhere in between.
Second, there was an overall effect of **context**, such that social context yielded more immediately rewarding both actions than the non-social context, consistent with our idea that the presence of an observer would facilitate focus on immediate rewards more in this particular situation.
Third, there was an interaction between context and goal such that there was an effect of context on action choice within learning and no-goal conditions but not for activation goal.

Results: Action choices



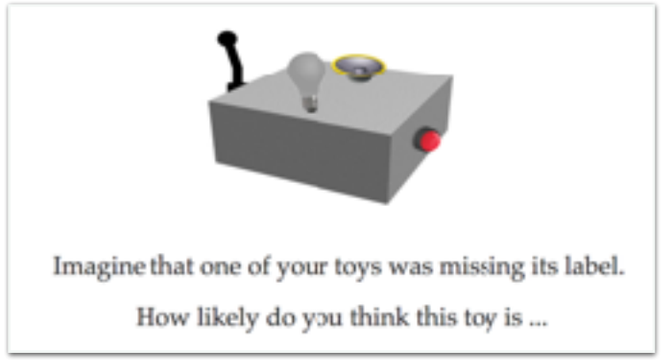
So here I'll show the results from participants' action choices. The X-axis shows the different goal conditions, and color represents social vs. non-social context. Y-axis shows the proportion of participants who chose the single action — so the action that is disambiguating but risking no immediate effect.
And here are the **results**. The first thing to notice is that participants **chose** more single actions in the order of learning condition, no-goal, activation, and presentation, which is consistent with what we had predicted. So having a learning goal leads participants to focus on more informative actions, whereas activation and presentation goals lead them to try to produce effects, and no-goal somewhere in between.
Second, there was an overall effect of **context**, such that social context yielded more immediately rewarding both actions than the non-social context, consistent with our idea that the presence of an observer would facilitate focus on immediate rewards more in this particular situation.
Third, there was an interaction between context and goal such that there was an effect of context on action choice within learning and no-goal conditions but not for activation goal.

Results: Information gain

To make sure that participants were actually correctly processing the informativity of single versus double actions, we also took a look at the amount of information gain based on participants' action choices in each of the goal conditions. We did this **by** collecting participants evaluation of likelihood for each of the possible hypotheses BEFORE and AFTER they made action choices and seeing their outcomes. If they are correctly processing the information from their actions and resulting outcomes, then their information gain should be higher for single actions than both actions since single actions disambiguate but both actions don't.

Results: Information gain

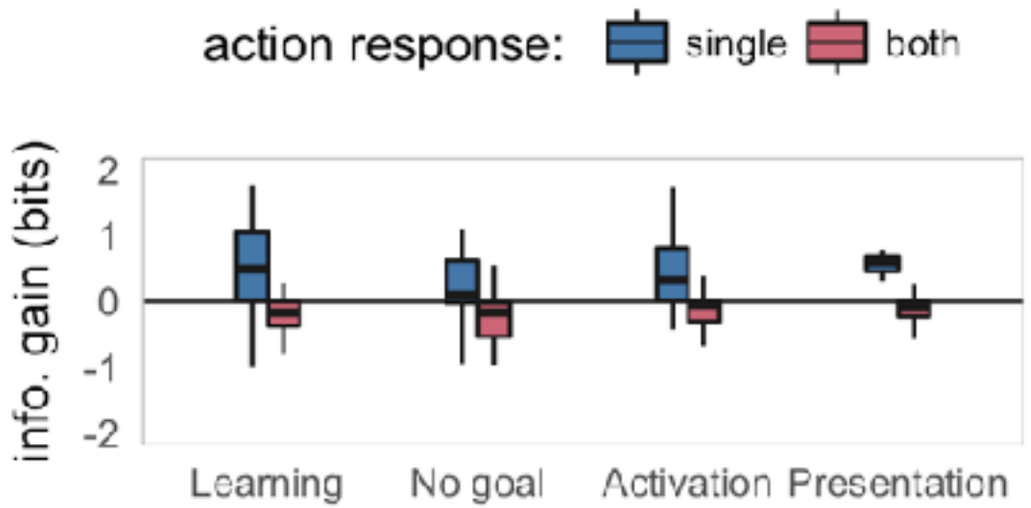
Before



After

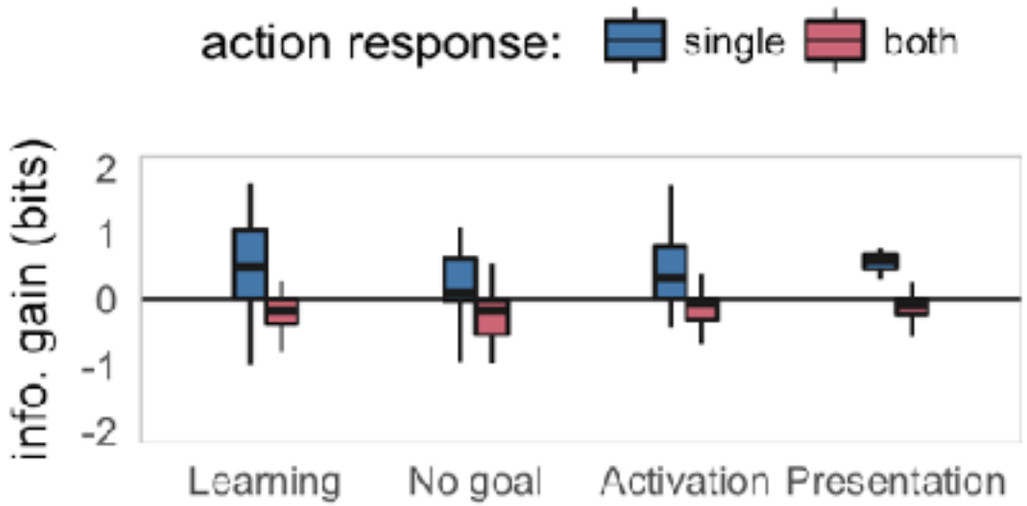
To make sure that participants were actually correctly processing the informativity of single versus double actions, we also took a look at the amount of information gain based on participants' action choices in each of the goal conditions. We did this **by** collecting participants evaluation of likelihood for each of the possible hypotheses BEFORE and AFTER they made action choices and seeing their outcomes. If they are correctly processing the information from their actions and resulting outcomes, then their information gain should be higher for single actions than both actions since single actions disambiguate but both actions don't.

Results: Information gain



Indeed, across all goal conditions, we saw that information gain was **higher** for single actions (in blue) compared to both actions (in red).
(and there was no difference between social vs. no-social context so that's not shown here).

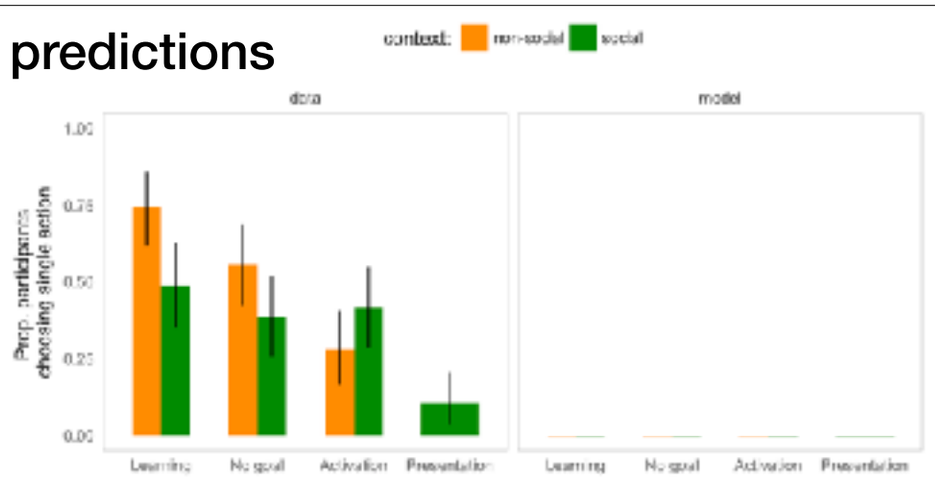
Results: Information gain



Information gain: single action > both action

Indeed, across all goal conditions, we saw that information gain was **higher** for single actions (in blue) compared to both actions (in red).
(and there was no difference between social vs. no-social context so that's not shown here).

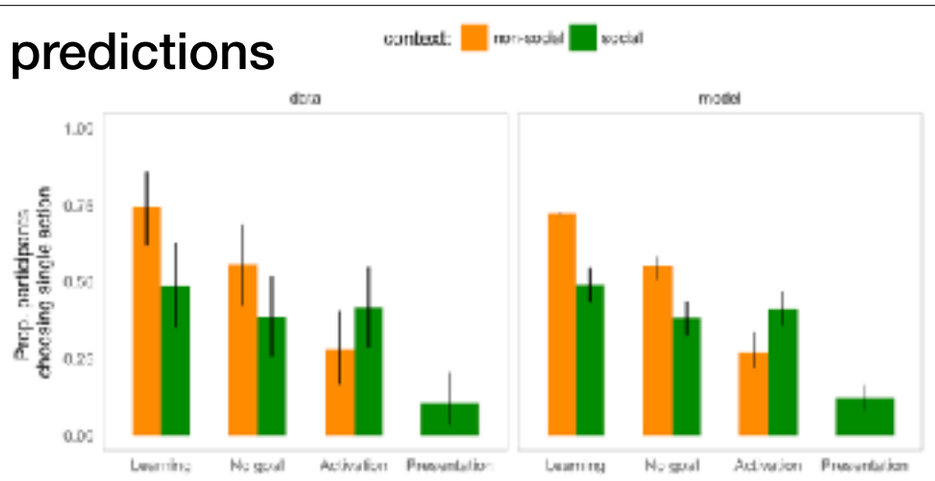
Model predictions



Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

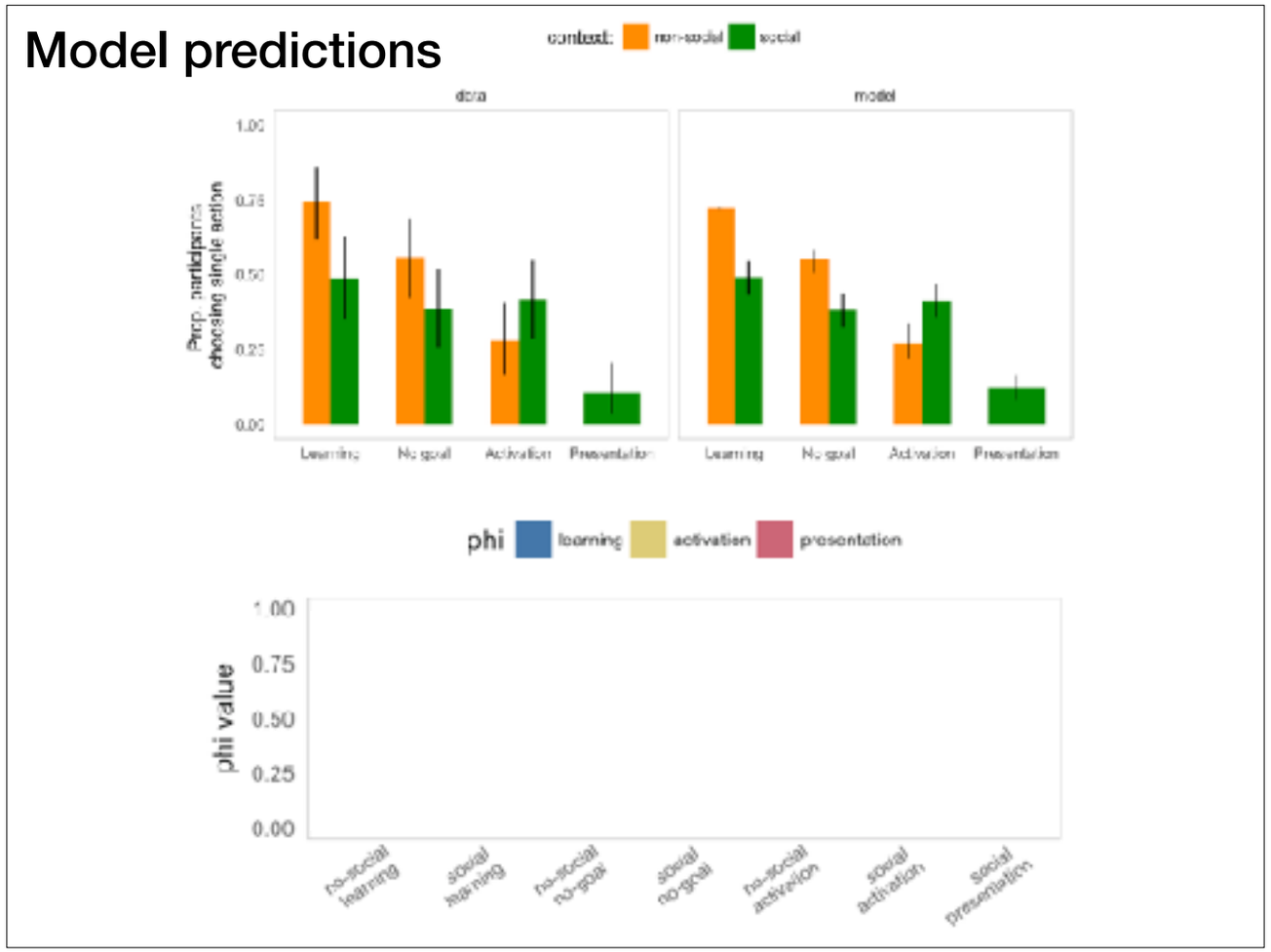
Model predictions



Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

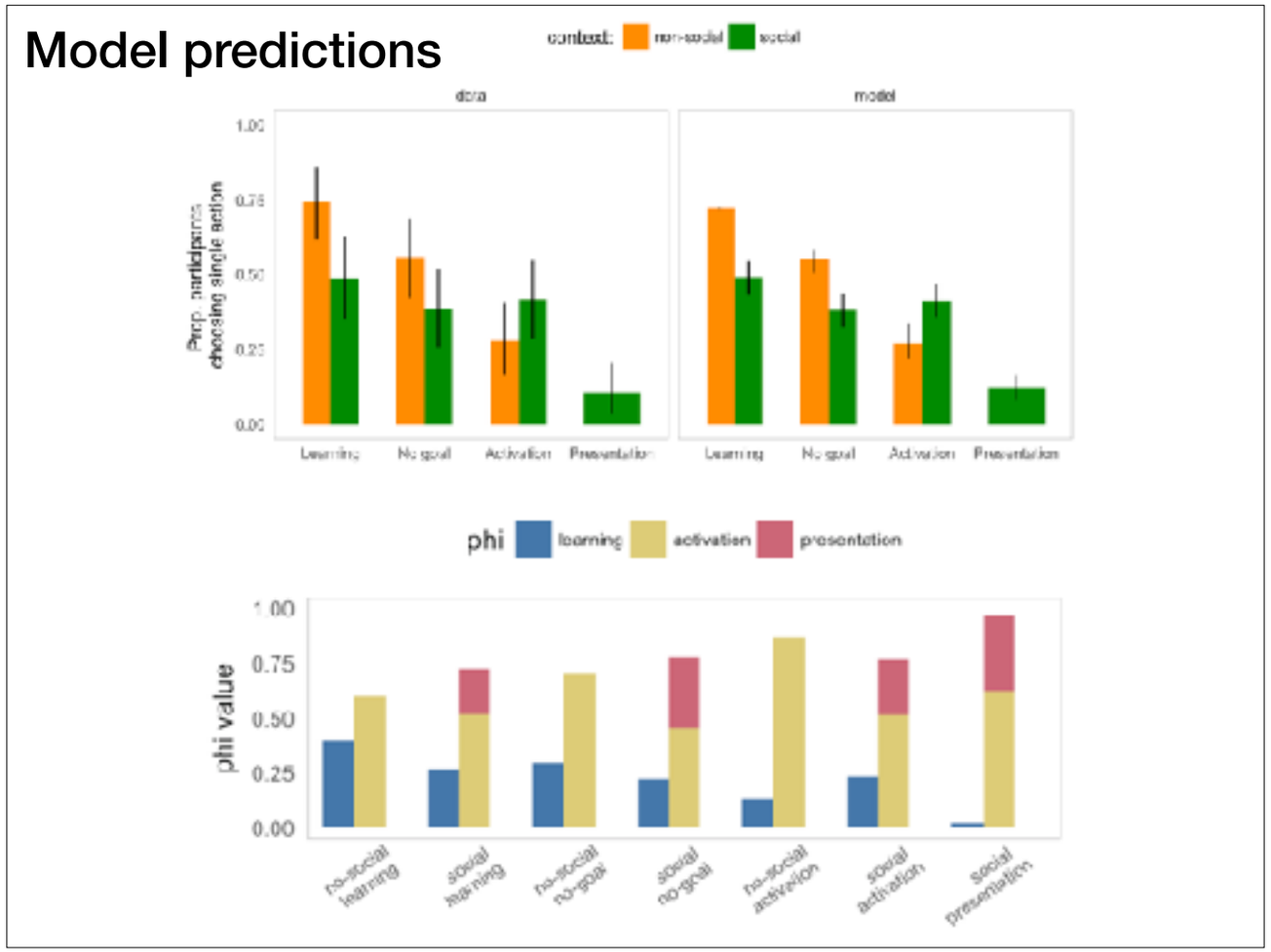
Model predictions



Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

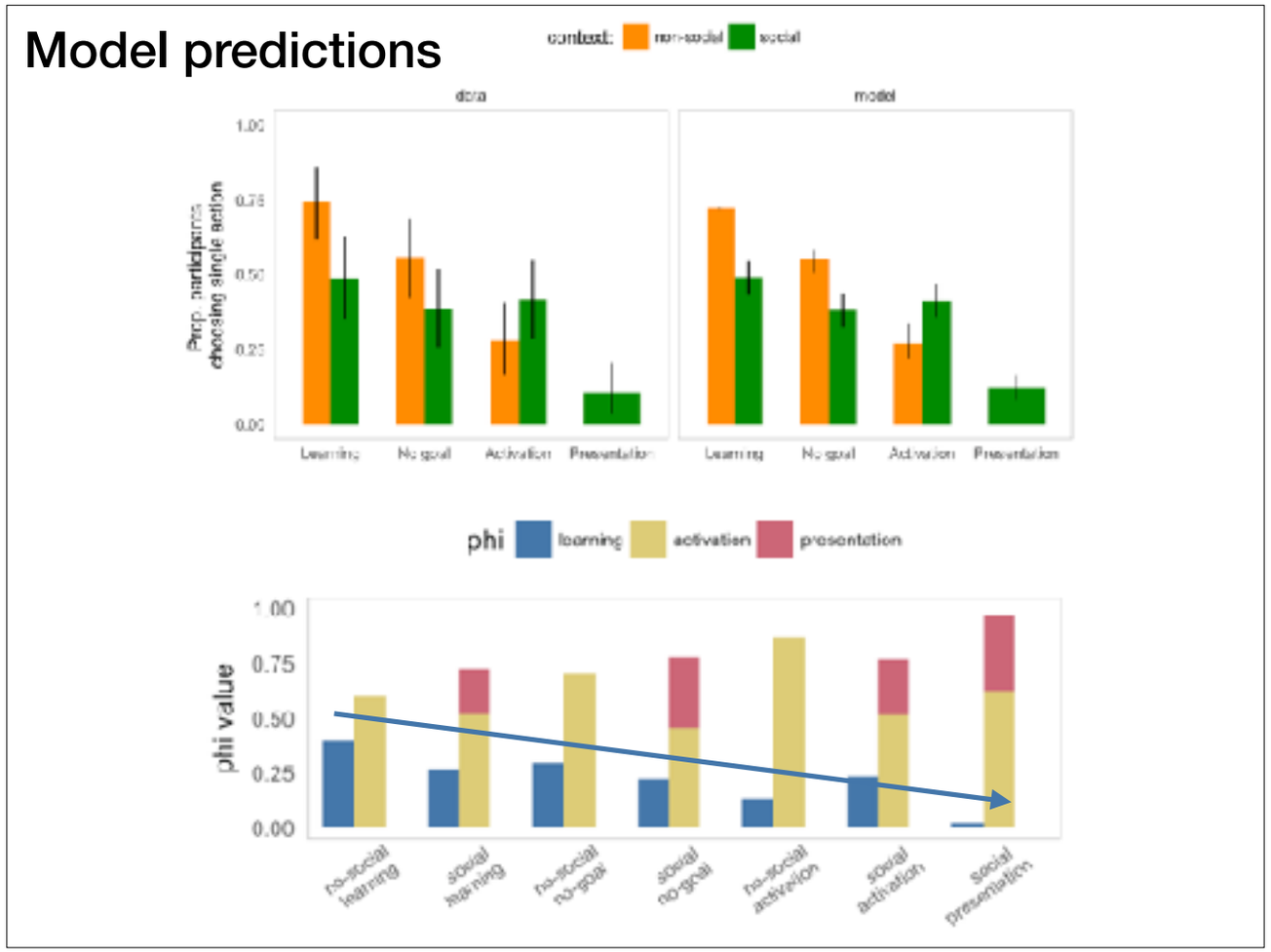
Model predictions



Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

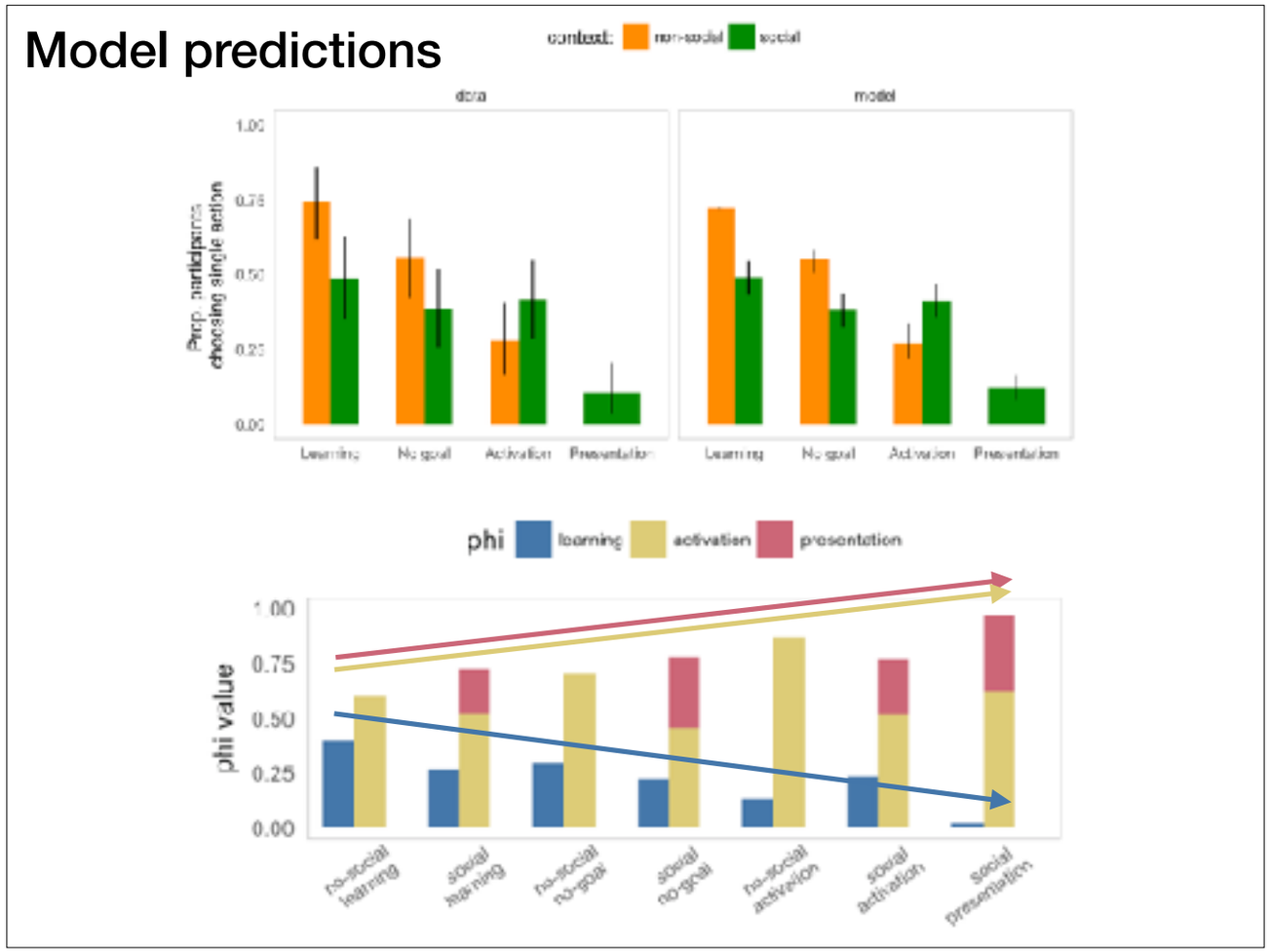
Model predictions



Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

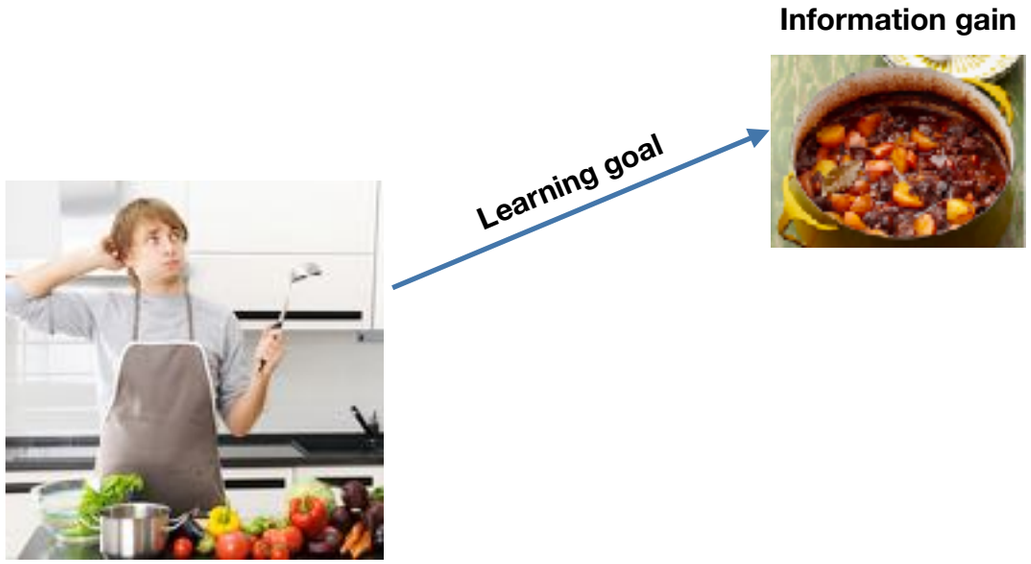
Model predictions



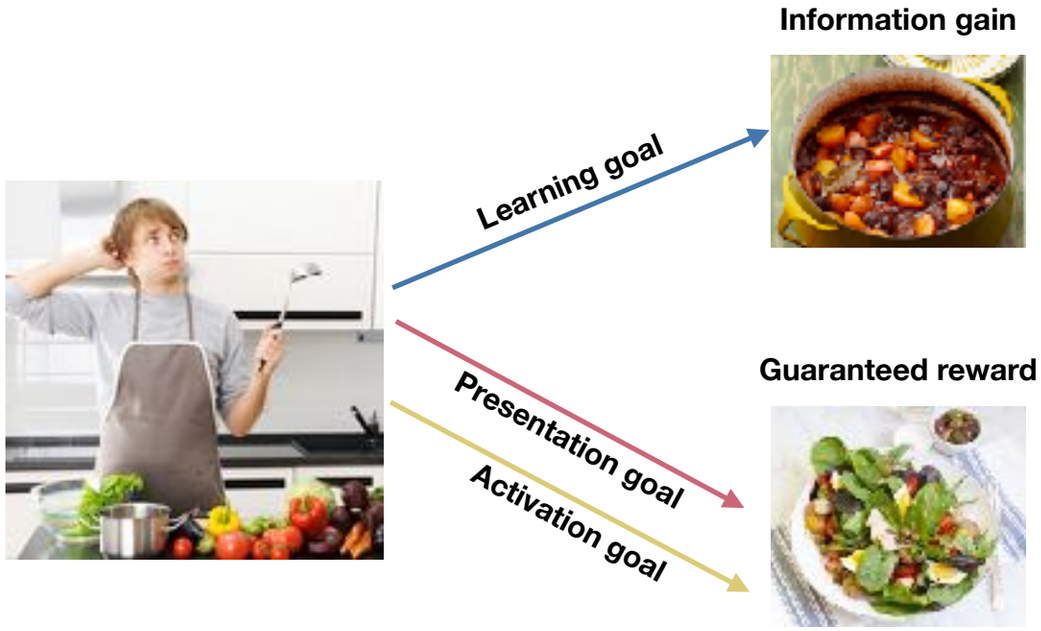
Finally, what did we see in the model predictions? The model **predicted** judgments very close to the human judgments — but it should be noted that this is because we let the model infer the weights for the utility for each context and goal combinations freely, so it's important to see how it predicted the **utility** weights. In other words, which utility or goal did the learner prioritize in each condition?

Here, the different colors represent the different utilities (blue is learning, yellow activation and red presentation) and x-axis shows the different conditions with context and goal descriptions. Y-axis is the inferred value of phi, which is the utility weight that shows the relative priority of the utilities in a given condition.
Here, we see that as we move from learning to no-goal to activation to presentation, we see **decreasing** weight on learning utility and increasing weight on activation plus presentation utility.
So these inferred weights were consistent with our prediction that learning utility would be highest in condition with learning goal and in absence of an observer, and that activation and presentation utility would be highest in condition with high social pressure to appear competent to your boss.

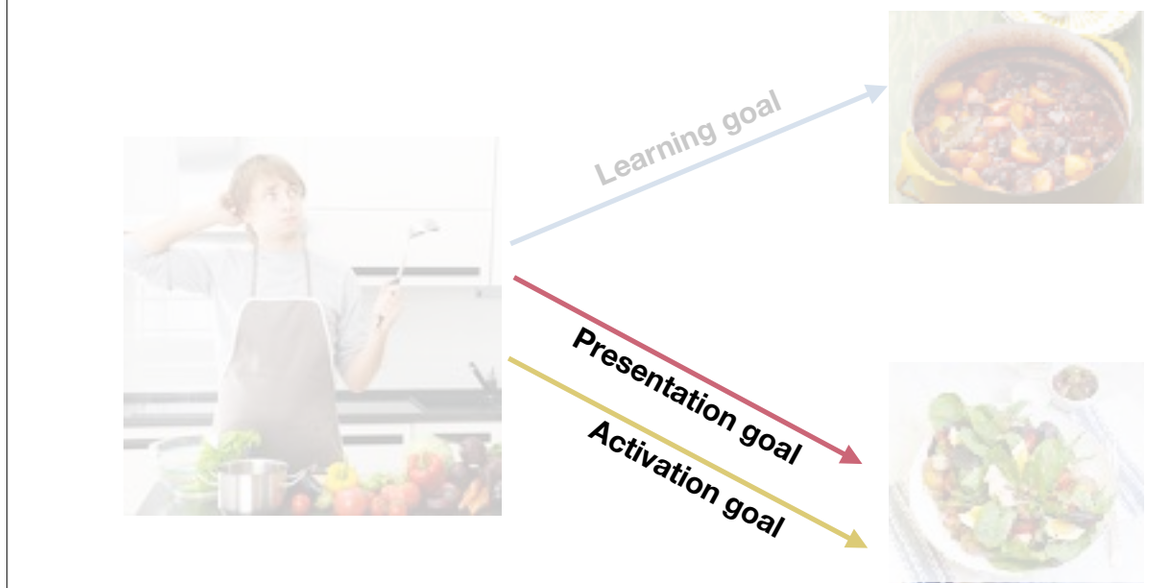
These results then might be able to generalize to our cooking example before: when learning goals are highlighted, people make choices that have higher risk but greater information gain, but when activation and presentational goals are highlighted, people choose lower-risk options that will guarantee a greater immediate reward.
But the current work is really a first step in this line of work and there are a ton of other things that should figured out to further apply our work to these kinds of real world situations.



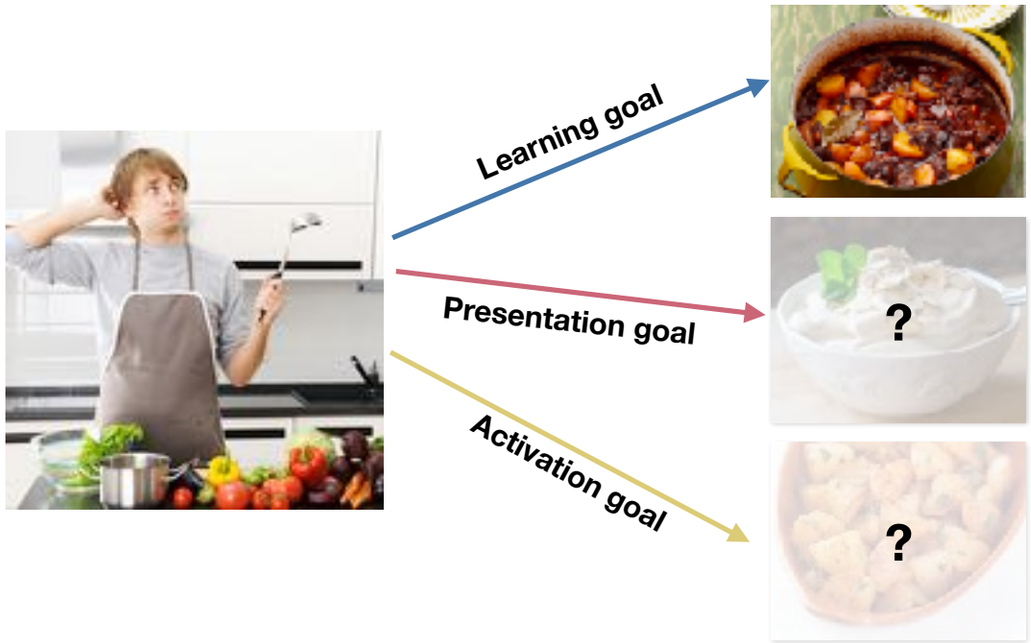
These results then might be able to generalize to our cooking example before: when learning goals are highlighted, people make choices that have higher risk but greater information gain, but when activation and presentational goals are highlighted, people choose lower-risk options that will guarantee a greater immediate reward.
But the current work is really a first step in this line of work and there are a ton of other things that should figured out to further apply our work to these kinds of real world situations.



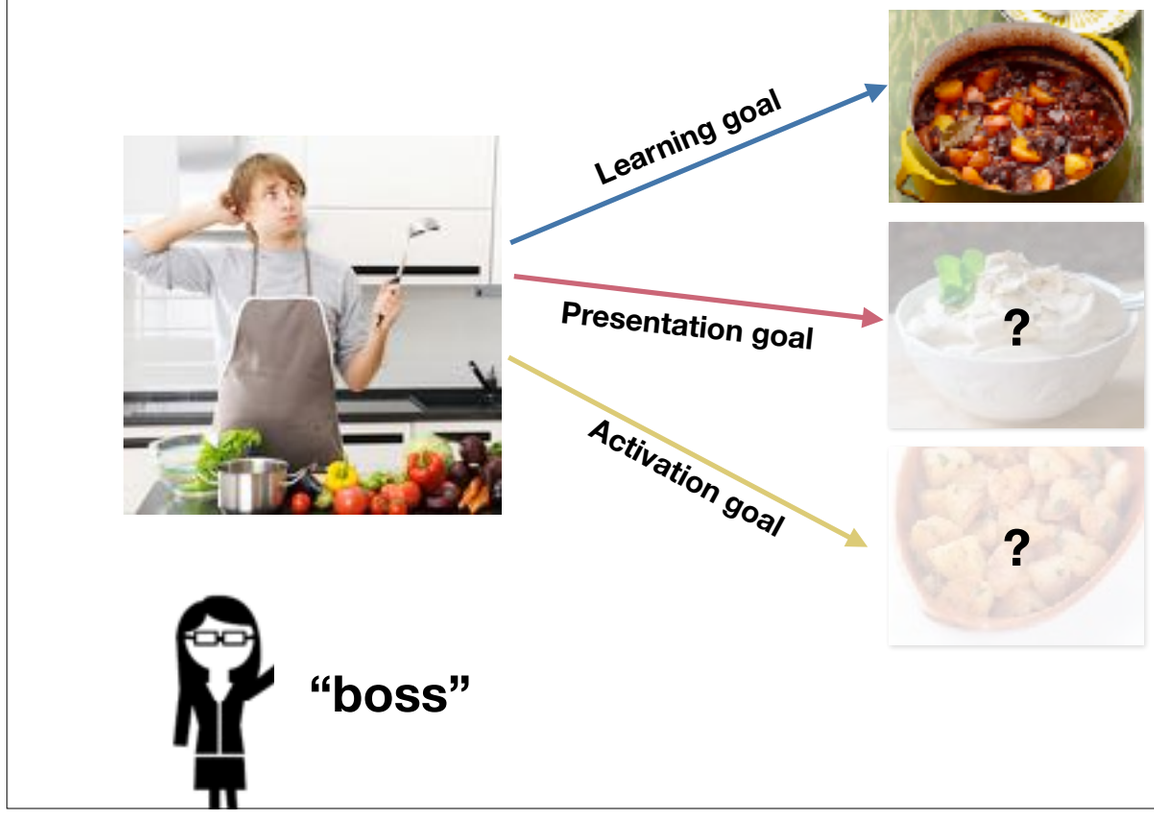
These results then might be able to generalize to our cooking example before: when learning goals are highlighted, people make choices that have higher risk but greater information gain, but when activation and presentational goals are highlighted, people choose lower-risk options that will guarantee a greater immediate reward.
But the current work is really a first step in this line of work and there are a ton of other things that should figured out to further apply our work to these kinds of real world situations.



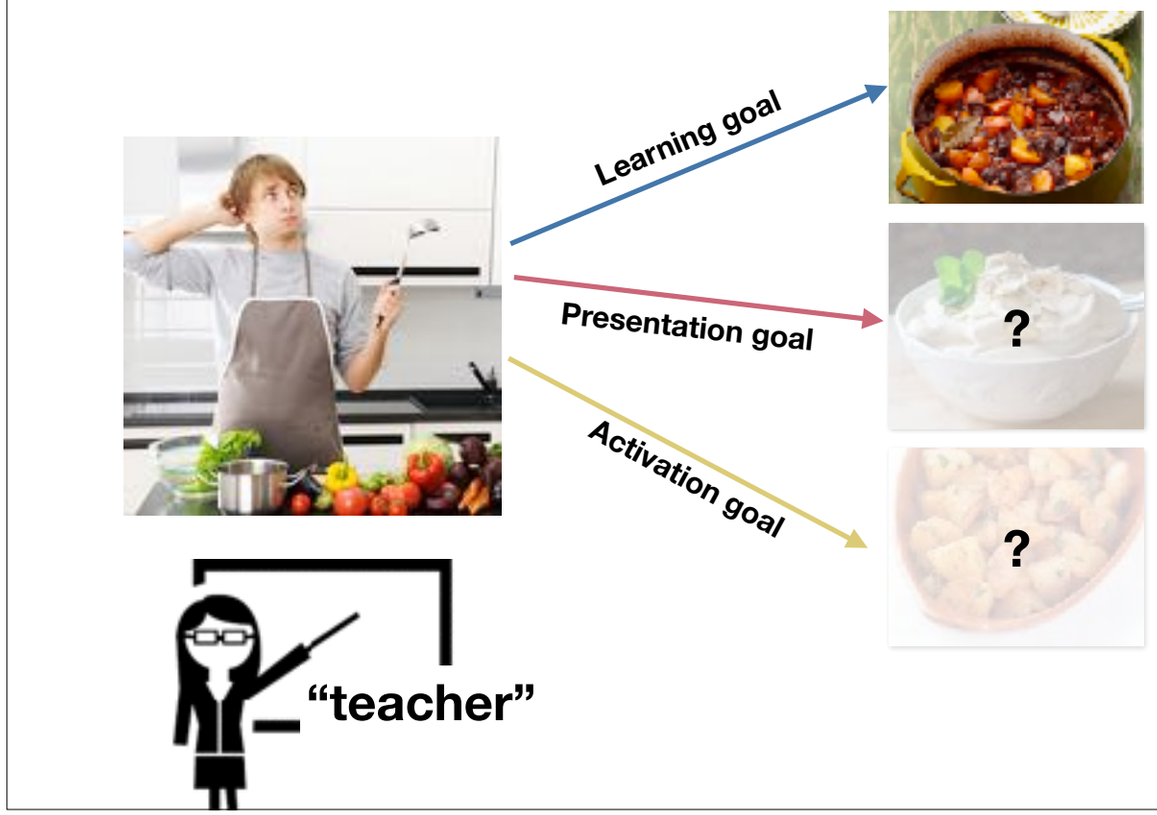
First, in our current model and paradigm, the activation and presentation utilities are not differentiated, as the choice of performing both actions satisfies both of these goals in our current task.



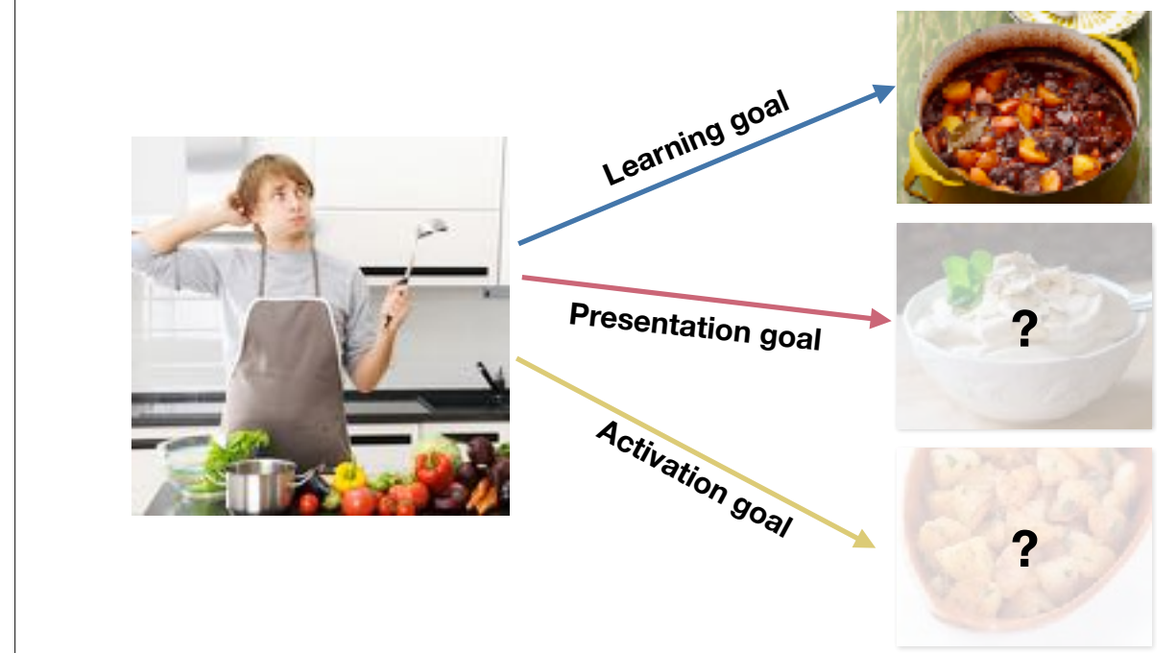
Enriching the space of possible actions could tease apart actions driven by self-presentation, especially when the optimal action for demonstrating one's competence may be different from the action for immediately rewarding outcomes.



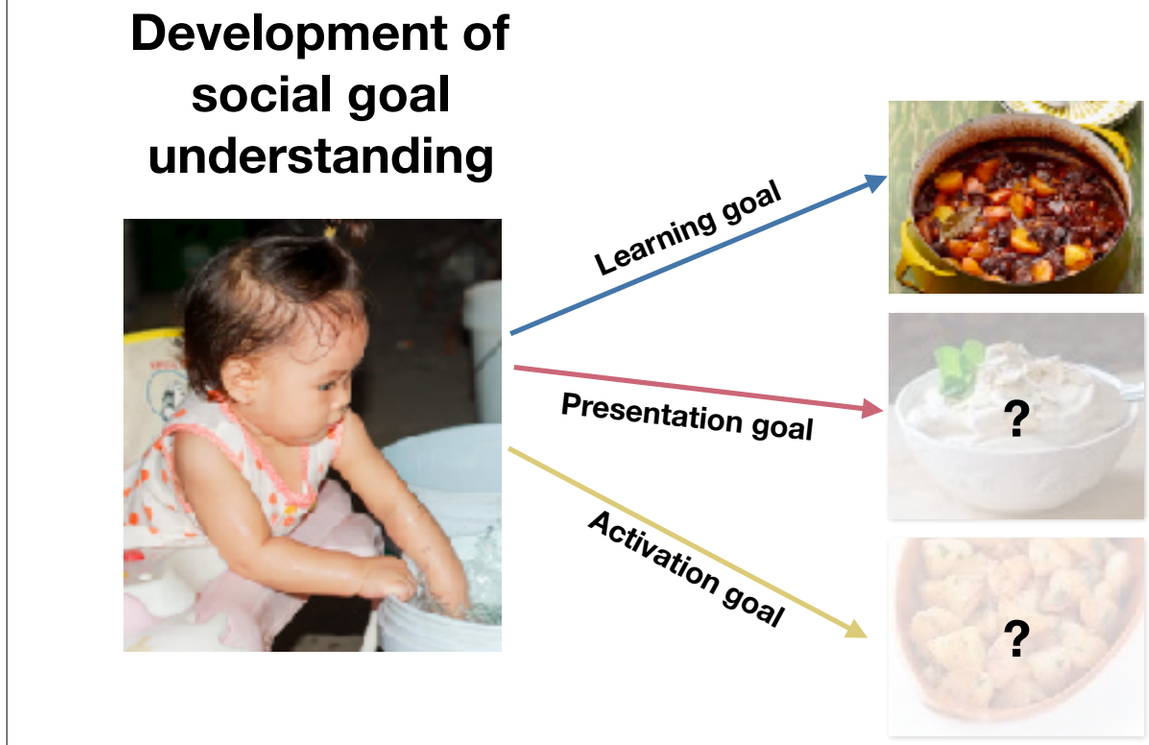
Second, we used a particular social context of the presence of a boss, but we can imagine many different kinds of social contexts that can influence the learner’s decisions very differently. For example, how would a learner behave in the presence of a **teacher**? In that case, the learner might not only think about appearing good to the teacher, but also consider the teacher’s goal for the learner to learn effectively, and thus may adjust his actions accordingly.



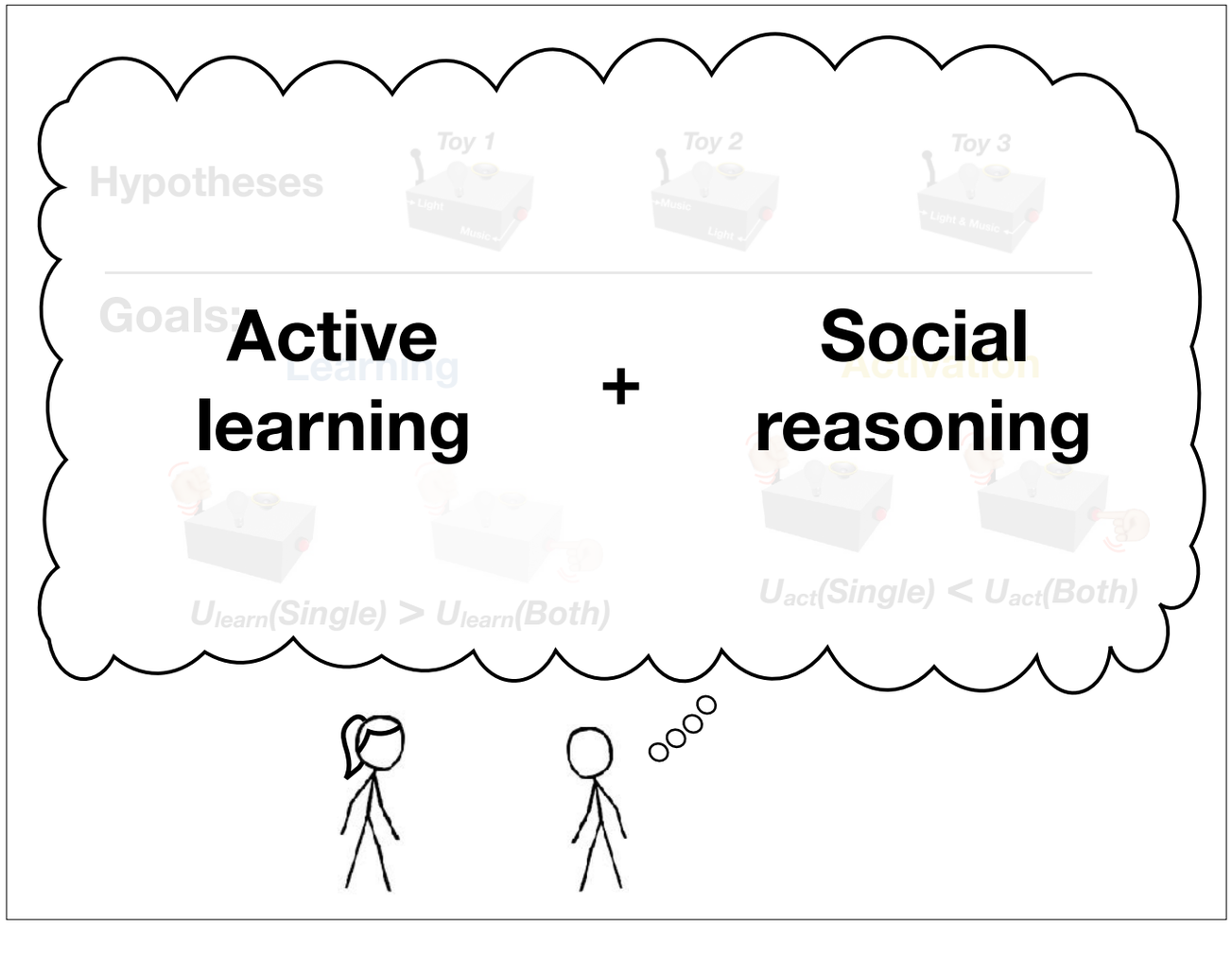
Second, we used a particular social context of the presence of a boss, but we can imagine many different kinds of social contexts that can influence the learner’s decisions very differently. For example, how would a learner behave in the presence of a **teacher**? In that case, the learner might not only think about appearing good to the teacher, but also consider the teacher’s goal for the learner to learn effectively, and thus may adjust his actions accordingly.



Finally, an interesting open question is how our model could be used to look at what **children** might think about this goal tradeoff. Do children also think about whether they look good when they explore? One possibility is that children may start out focusing on learning goals when surrounded by familiar caregivers who scaffold learning-relevant actions; but as their social abilities mature and their social environments become more complex, children might start to emphasize their presentation goals much more.



Finally, an interesting open question is how our model could be used to look at what **children** might think about this goal tradeoff. Do children also think about whether they look good when they explore? One possibility is that children may start out focusing on learning goals when surrounded by familiar caregivers who scaffold learning-relevant actions; but as their social abilities mature and their social environments become more complex, children might start to emphasize their presentation goals much more.



Overall, we believe that this work represents a first step to answering these rich questions that are relevant to our learning experiences that so often occur with people around us, and ultimately seek to bring together theories of active learning and social reasoning.

Thanks!

All experiments, data, model, and analysis codes are available in the public repository for the project: https://github.com/kemacdonald/soc-info