Supplementary Material

A Model Inputs

For extra clarity, we show here the exact input representation given to our models when including all the grounding features we consider in the experiments (setting, objects, characters + personas, actions, emotes, and dialogue). An example is given in Figure 2.

We note that there are other ways to represent this information that we have not explored that could improve performance. Further, there is additional information in LIGHT that could possibly be encoded in the input text: for example, what characters are carrying, and the affordances of objects. The latter, while not explicitly provided in the input does constrain the available actions, so it is still used by the model. Object affordances such as *is gettable* are visible to models via the action history, but more explicit inputs could potentially be useful, and this could be explored in future work.

B Bi-Ranker and Cross-Ranker Speeds

We give test time computation speeds for the BERT-based Bi-Ranker and Cross-Rankers in Tables 10 and 11 for the emote and dialogue tasks. For the emote task, the Cross-Ranker is still feasible due to there being only 22 labels to compute, although it is still 4.6x slower than the Bi-Ranker if the 22 candidate representations are cached. The Bi-Ranker can always cache label representations if they are fixed for many input examples (the common case) because the representation does not depend on the input. For the Cross-Ranker this cannot be done because the label representations are contextually dependent on the input. For dialogue retrieval, because the number of candidates is so large (more than 100,000) caching makes the Bi-Ranker feasible whereas the Cross-Ranker, which cannot cache label representations, is infeasible to compute.

C Unseen Test Set Overlap

The unseen test set is chosen by design to be relatively distinct from those available in the training set, and the actual content (descriptions, personas, dialogues) are entirely disjoint. However, due to the large size of the dataset, it is possible the names of locations, characters, and objects in

Emote	Bi-Ranker	Cross-Ranker
w/o caching	171s	326s (~1.9x slower)
with caching	70s	n/a (~4.6x slower)

Table 10: Bert Bi-Ranker and Cross-Ranker speeds on the emote task, test seen (2495 examples), 22 candidates per example.

	Bi-Ranker	Cross-Ranker
Dialogue	2.07s	24453s (~11812x slower)

Table 11: Bert Bi-Ranker and Cross-Ranker speeds on the dialogue task, per single example average (retrieval over 110,877 training set candidates).

the unseen set could have word overlap. We assert this by comparing word overlap with the names of locations, characters, and objects in the training set. Of the 73 locations, 207 characters, and 956 objects created from the unseen location categories, the names of 3 locations, 96 characters, and 203 objects exactly match names of elements in the training set. We note that these represent names such as *tavern*, but the chats are collected with the full location descriptions (which are unseen in the training set) and thus reduces overlap with train.

D Crowdsourcing Methodology

Expanding on the dataset collection explanations in section 3, a number of steps were taken to attain a level of quality and consistency. The first and most influential came from the constrains of the setting itself. We used a fantasy setting to try to encourage some kind of continuity across the dataset. We believed that workers would share some kind of common understanding about what a fantasy environment would entail, and then this understanding would be reflected in the dataset. It also ensured there were easy ways to flag certain workers that were creating content that wouldn't make sense in the dataset (referencing real locations, modern day objects, etc.). From here we could remove some content and filter workers out from continuing to work on this dataset. The other primary technique regarded using rounds of pilots and staged tasks to gradually filter towards high quality content rather than collecting all of the content in a single forward pass. Nearly half of the content in each initial pilot task was discarded, and we iterated on pilot tasks until the discard rate was

Input to Model:
_task_speech
_setting_name main foyer, Inside Castle
_setting_desc The main foyer is massive. A grand staircase sits to the back of the foyer leading to the upstairs.
At the front of the foyer stand two servants ready to help anyone who comes to visit. To the left of the room there
is a doorway leading into a corridor. To the right there is a door leading to another corridor for the King's servants.
At the foot of the stairs there is a bearskin rug that is staring at you almost as if still hungry. The walls are
lined with portraits of the king and his family.
_partner_name servant
_self_name king
_self_persona I am a king of the whole empire. I give rules and pursuit them. I am brave and fearless.
_object_desc a duster : The duster has large gray feathers bound together by a leather wrap.
_object_desc a small bucket : The bucket may be small but it gets the job done.
_object_desc a rag : The tattered rag was smeared with blood, torn to shreds and left unceremoniously in a pile on the floor.
_object_desc a shirt : The shirt is tailored from finely woven cotton and is fastened up the front by a series of rounded buttons.
_object_desc a crown : Thought of as a holy item, the crown goes only to those who are worthy enough.
_object_desc a scepter : On its handle, you see two red gems gleaming like eyes of an animal.
_partner_say my humble king. What am I to do to serve you?
_self_act give scepter to servant
_partner_say Yes my lord. I will polish it immediately. Am I to return it to you personally?
_partner_act put scepter in small bucket
_self_act give crown to servant

Label: Yes. Yes. Of course. Also check the jewels in my crown. They seem loose.

Figure 2: Example input format (and target label) given to models, following the same dialogue in Figure 1. Tokens like "_setting_name" are special tokens intended to be signifiers for the encoding module of a network to know which piece of grounding information is being read on that line.

less than 1 in 30 tasks. The rest of this section will discuss some specific measures taken at the individual task level, and will acknowledge some arguable deficiencies and potential areas of improvement on the dataset in its current form.

Locations The location task of creating a description, backstory, list of connected rooms, and annotations of characters and objects present seemed to be too disjoint of a task based on the crowdsourcing best practice of breaking down tasks into as atomic of an action as possible. Thus we split it into two tasks, the first to provide the core text content and list of connected rooms, and the second to annotate the content inside those rooms. We will refer to these as *Task 1* and *Task 2*, and were simple form-entry tasks as displayed in Figures 4 and 5. These two tasks were used in sequence to produce the locations present in the dataset.

In order to drive quality, we manually reviewed a handful of rooms from each worker to assert that the rooms had proper English descriptions and back-stories, and that the room fit appropriately in the category provided. In retrospect, given the two-tiered task setup and some of the techniques we developed later in the collection setup, we could have asked workers who were annotating rooms in Task 2 to provide some kind of signal about the quality of the rooms from Task 1 in

order to have a lower-cost method for evaluating the quality of the work from Task 1 than using our own time.

Ultimately, one of the most important steps for improving dataset quality at this stage was creating form validators that caught the most common error cases from the first time around. These validators had the bonus effect of deterring botting of our tasks, as they couldn't pass the validation stage. For Task 1, the simple validator we ended up using asserted at least one complete sentence (determined via capitalization and punctuation) for both the description and background. For Task 2, our validation step forced workers to enter values that had direct word overlap with the entered text.

One of the largest difficulties with Task 2 was that some workers would optimize for grabbing key words out of the text without taking the time to fully understand the context. As thus, phrases like *"and the remains of adventurers long dead"* would occasionally result in workers annotating the presence of *adventurers* as characters in the given room. We attempted to mitigate this type of false positive with both explanatory examples and spot checks to soft-block workers who made this mistake consistently. At the moment a small number of these still remain in the dataset, but generally in instances where it still makes sense as in

the above example, where the room definitely has remains of previous adventurers, but appropriately could also have some current adventurers as well.

Characters Similarly to how we split Location collection into two tasks, Character collection was split into two tasks as well. The first asked workers to clean up the span selected in Task 2 in order to remove words that didn't directly relate to or describe the character, and to provide a singular form for plural characters (as we intended for someone to eventually play the role of the singular character), tag the character as a person, creature, or object that was accidentally tagged as a character, and then asked for a *first-person* perspective persona for the singular character. The second task gave workers the name of a character and their persona, and asked for a *second-person* perspective description for the character as well as a list of objects that the character may be carrying, wielding, or wearing. We'll call these tasks *Task 3* and *Task 4*, and these were also collected via form-based tasks as displayed in Figures 6 and 7. We used complete sentence form validation for both the persona from Task 3 and text descriptions in Task 4 to flag potential bad examples to filter out.

The goal of the Task 3 was two-fold, first to validate and standardize the format of output from Task 2, and then second to begin to collect the creative content in the form of a persona. For example, we used Task 3 to transition from *Sneaky Thieves who stole the gold* to *Sneaky Thieves* to *Sneaky Thief*. Based on worker feedback from initial pilots, we found that balancing creative and mechanical work in the same task kept workers more engaged with the tasks at hand.

The most common mistake that surfaced in the initial pilots was incomplete entries for tasks that didn't actually require correction, for example if the provided form was simply *Traveler*. We chose to embrace this format and assume that unfilled entries were already in their base form. The second most common mistake was describing personas from a third person perspective. This occurrence required manual filtering, as in some cases it was actually somewhat character appropriate to have a persona in that format, such as for an uneducated goblin. We filtered out a majority of these by searching for direct overlap between the provided character name and the persona. Ultimately it's easy to extract the examples that have

the clearest grounding format by filtering for examples that contain *"I"*, so as these examples provide more variety in the dataset we chose to keep them.

A remaining issue brought forth by our singular-form constraint is that it was somewhat ambiguous how one would get the singular form of a collective term such as *family*. In most cases we found that workers would choose to provide the format of *collective member* or simply *person*, which sometimes led to vague personas and thus less strong grounding in followup tasks. The content is still workable in these cases though, just not as ideal as we might have wanted. A possible route for improvement here would be a task that asks workers to create a few possible members for a collective for any character we currently have annotated as a member. It is important to note that these cases account for just 44 out of the 1755 collected characters.

One issue of note that surfaced in Task 4 was that workers occasionally described clothing that would potentially lead to risky actions and conversation material, so we chose to eliminate undergarments from the dataset to prevent the creation of inappropriate combinations with the *remove* action. This was included as something to not write about in the task text.

Objects The object task is most similar to Task 3, but refocused on annotating objects that were specified in Tasks 2 and 4. It took a step to correct the provided span and give a textual description of the object. It also asked for a number of affordances, namely if the object can be picked up, is a container, is a surface, can be eaten, can be drank, can be worn, or can be wielded. We also collected a flag for if a particular example was not appropriate for the dataset or was hard to make sense of. This content was also collected as a form-based task, and we refer to it as *Task 5* and display it in Figure 8. We use complete sentence validation on the text description as a simple quality filter as in previous tasks.

The methodology for Task 5 is very similar to Task 3, trying to both standardize data from previous tasks and act as a filter for bad content that could have been overlooked before. It similarly had both a mechanical data entry and creative component, which tried to keep engagement up.

Overall the largest problem that was surfaced in the pilots was that workers tended to come up

with descriptions for objects that were incompatible with our long term goal of having modular components that can be mixed and matched between rooms and scenarios. This came up in many forms, such as workers describing objects as if they were being used in a scene happening in the present, as in *the sword glimmered in the hands of the knight, wielded high in the sky in a call to battle*. While creative, these ultimately were not what we were looking for, so we explicitly called out descriptions like this and many others as being undesired content in our task description. We then manually checked a few examples from each worker to ensure that the data coming in for the final task mostly adhered to this rule.

It is important to note that the object affordances collected are somewhat noisy due to different possible interpretations of the primary object or the tags. Something like a *branch* could be valid as a surface in one circumstance, or a gettable weapon in another. We attempted to reconcile some individual affordances where the pairings of affordances didn't make much sense (for example, very few objects should be both a weapon and edible). This helped with certain objects that were over-tagged, however we haven't used any methods for reconciling scenarios where an object was under-tagged.

Dialogues   Dialogue collection was the hardest task to get correct, and required the largest number of pilot tasks and worker quality control techniques to get to a place that we were satisfied with. The final approach included creating a simple but deliberate onboarding test that needed to be passed in order to move forward with the task at all, collecting mutual feedback from workers about each other, setting timeouts for how quickly workers needed to respond to each turn, and manually validating a few examples from each worker. Each of these steps aimed to solve a different problem, as described in the rest of this section. We will refer to this task as *Task 6*, and it was collected using the ParlAI-MTurk interface as shown in Figure 9.

Firstly, we needed to pair two workers together in order to properly collect dialogues with people playing two different roles without necessarily having insider information into the decisions of each others' turns. While pairing workers solves this problem, it makes the worker experience incredibly dependent on the quality of the worker that they are paired with you. Furthermore, if a worker is paired with a worker that is extremely low quality, the whole dialogue may need to be discarded or is otherwise only useful as an example for how a model might want to react to bad input. If the other worker is good, this makes having any bad workers in the pool not just a poor experience for workers but expensive for the collection process in general. This is the problem that the initial onboarding test aimed to solve. The requirements for passing included entering a specific correct answer as well as at least 4 characters of into the text field. The required action was created such that a worker would have to read and understand the provided persona and setting, how the two interact, the characters and actions available, and be able to synthesize all of the information with an understanding of how to use the interface to send the correct answer. The test required getting the single action correct in 3 attempts. Failing the test on any attempt would permanently soft block a worker from working on Task 6 in the future.

The above test did a lot of work for flagging workers that were well below the bar for completing Task 6 at the level we wanted for the dataset, however as it was a one turn test and it had no way to fully evaluate the quality by which workers would actually incorporate their persona and the setting into their dialogue turns. Furthermore, it didn't filter out workers that would take too much time on their turns and thus cause their partners to disengage and provide lower quality responses, potentially due to working on other tasks in the background and doing too much context switching. We solved these problems separately.

In order to handle low quality workers, we allowed workers the opportunity to rate each other at the end of each dialogue, and to provide tags about the experience. We found that positive feedback was generally noisy and hard to get signal from, but negative feedback almost always correlated to a worker who was providing bad content. As a bonus, workers gave us positive feedback about this capability, as it allowed them to filter out workers that made the task less engaging and interesting for them. We reviewed this feedback periodically while tasks were running and soft-blocked workers low quality workers whenever they were flagged.

In order to handle the influence of response time on task quality, we set a maximum response time of 5 minutes for any given turn, and overall started

soft blocking workers that were consistently above 2 minutes for each message, even if their particular content was pretty good. This improved collection times and did not seem to negatively affect quality.

After this point, manually checking the collected conversations still surfaced a few bad examples when viewing one chat per worker rather than arbitrarily sampling the dataset. In order to remedy this, the last quality check was a direct evaluation of at least 2 dialogues from each worker. This caught a few overlooked instances from workers that didn't necessarily work on enough tasks to get flagged by one of our consistently reviewing workers. Generally this surfaced some quality issues surrounding profanity, inappropriate content for the given setting, and entire misunderstanding of the task at hand such as never using the persona or location as grounding context in the conversation. As not all workers were particularly diligent raters (as confirmed by the low signal of positive ratings - workers don't necessarily want to flag each other as bad), a few workers were able to slip through the cracks up until this point due to not completing enough tasks to encounter a rater that flagged them.

One small acknowledgement throughout the dialogues is that there are still misspellings, improper grammar, mistaken keystrokes, and such. While the rate of occurrence is orders of magnitude lower than we observed in the initial pilots, it is hard to separate cases where it is a genuine mistake versus cases where it is appropriate for the character, such as a *pirate* using seaworthy lexicon and adding extra R's to suggest a pirate-like drawl, or a *snake* that slips in extra S's to better play the role.

E   Descriptions of Actions and Emotes

The LIGHT action set builds upon the graph framework introduced in Mastering the Dungeon (Yang et al., 2017). The basic idea presented is that everything in the text adventure game can be represented as nodes, and then state is described by edges between those nodes. In this way, an agent and an object can be in a room, and that agent can be carrying a different object or a container might have an object inside as well by the same kind of relation. After defining this relationship, we can further define a set of actions that can be taken based on a combination of the state of the graph and the attributes of nodes in that graph.

---

applaud, blush, cry, dance, frown, gasp, grin, groan, growl, laugh, nod, nudge, ponder, pout, scream, shrug, sigh, smile, stare, wave, wink, yawn

---

Figure 3: Emote options within the LIGHT platform

The available actions for the dialogues collected in this dataset, along with the constraints for applying those actions, are available in Table 12. We used the crowdsourced object affordances to set the correct attributes for nodes in the graph (if the object can be picked up, is a container, is a surface, can be eaten, can be drank, can be worn, or can be wielded).

For the emotes, we paired down a list of emotes sourced from existing MUDs to reduce redundancy and task complexity at the acknowledged cost of expressiveness. This led us to select just one out of *scream*, *shout*, and *yell* instead of keeping them all, as having all of the emotes would lead to a more complicated crowdsourcing task than we wanted to risk. We ended up with a set of 22 emotes, listed in Figure 3.

F   Descriptions of Human Evaluations

As crowdworkers can sometimes be inconsistent, we set up two filters to onboard workers into being fair representatives for human perfomance on the task. The first gave workers a few chances to select the correct input for a turn each of dialogue, emote, and action on a scenario we created to strongly hint at the correct answer. We then chose to use performance on the training set as a secondary filter to have workers that were capable of the task. Each of the tasks has a different level of difficulty, so we selected reasonable benchmark values based on our own performance on the tasks. For dialogue, this required getting all 7 of the turns from the training set correctly. For actions, this required getting 6 out of 8 turns from the training set correctly. Lastly for emoting, we required getting only 2 out of 8 turns from the training set correctly. On the seen set, our accuracy on the dialogue, action, and emote tasks were calculated from 217, 165, and 211 turns respectively. On the unseen set, we calculated the accuracy from 196, 114, and 209 turns respectively.

G   Embedding Visualizations

To explore the diversity of LIGHT, we use t-SNE (van der Maaten and Hinton, 2008) to visualize

Action	Constraints	Outcome
get *object*	actor and *object* in same room *object* is gettable	actor is carrying *object*
drop *object*	actor is carrying *object* *object* is gettable	*object* is in room
get *object1* from *object2*	Actor and *object2* in same room *object1* is gettable *object2* is surface or container *object2* is carrying *object1*	actor is carrying *object1*
put *object1* in/on *object2*	Actor and *object2* in same room *object2* is container or surface actor is carrying *object1*	*object2* is carrying *object1*
give *object* to *agent*	Actor and *agent* in same room *object* is a member of actor	*agent* is carrying *object*
steal *object* from *agent*	actor and *agent* in same room *object* is a member of *agent*	actor is carrying *object*
hit *agent*	Actor and *agent* in same room	inform *agent* of attack
hug *agent*	Actor and *agent* in same room	inform *agent* of hug
drink *object*	actor is carrying *object* *object* is a drink	inform actor of drinking successfully
eat *object*	actor is carrying *object* *object* is a food	inform actor of eating successfully
wear *object*	actor is carrying *object* *object* is wearable	actor is wearing *object*
wield *object*	actor is carrying *object* *object* is a weapon	actor is wielding *object*
remove *object*	actor is wearing/wielding *object* *object* is wearable or a weapon	actor is carrying *object*

Table 12: LIGHT actions and constraints

the embeddings of the different atomic dataset elements – locations, objects, characters, and actions. We use two different embeddings methods to tease out two key aspects of our dataset: 1) the *interconnectedness* of grounding information (relationships between different types of elements, such as the actions available around given objects, or in a given location), and 2) *coverage* (the variety of different objects, locations, and characters in our world).

To explore the *interconnectedness* of our dataset, we visualize the embeddings learned when training the baseline Starspace ranking model on the task of dialogue, action, and emote prediction, in this case with no pretrained vectors so learning comes from our dataset alone. The t-SNE visualizations of these Starspace embedding can be found in Figure 17. Because the Starspace model operates by mapping all inputs and outputs to a shared embedding space, we find the learned embeddings capture many of the nuances and relationships between different elements

of our dataset. For example, looking at the nearest neighbors for the location "Dock" (the bottom-right of Figure 17), we see actions like "get crate from ship," "put plank in ship," objects like "ship" and "rope," and characters like "boat workers." We see similar relationships captured when looking at nearest neighbors for the "painters" characters, the "hug horse" action, and the "pillows" objects.

To explore the *coverage* of our dataset, we use pretrained GLoVe word embeddings (Pennington et al., 2014), trained on the Common Crawl corpus. As each dataset element can consist of multiple words (e.g. "give the horse a potato," or "The Queen's Chamber"), we take the mean of the GLoVE vectors for each word as the fixed vector embedding for the element. The t-SNE visualizations of these GLoVe-embedded elements can be found in Figure 18. Unlike the Starspace embeddings, which capture the structure present in the relationships between different types of dataset elements, we find that the GLoVe embed-

dings capture the breadth and semantic similarities of dataset elements. For example, looking at the nearest neighbors for the embedding of the "Dock" location, we see similar locations present in our dataset, like "Ferry Terminal," "Wharf," "pier," and "Boathouse." Similarly, if we look at the nearest neighbors for the "pillows" objects, we see other objects like "bedding," "mattresses," "rugs," "towels," and "curtains."

H Action and Emote Relationships

To visualize the interaction trends between actions and emotes in LIGHT, we present heatmaps (in Figure 19) counting the number of occurrences of each immediately before or after one's partner performs an action or emote. While responses to an action or emote can evolve over multiple timesteps, we limit this visualization to action relationships within a single timestep. Additionally, to effectively measure trends in physical actions, we cluster all physical actions by the root word (for example, "steal the sword from the soldier" becomes "steal").

While for the most part there are a multitude of different observed physical and emotional responses for each partner move, there are certain interesting trends to observe. Looking at the top-left of Figure 19, we see that if one's partner makes a "hit" action, the most likely response is to "hit" back. Looking at the same plot, we see that "hug" actions are similarly reciprocated. If we look at the interplay between physical actions and emotes (top-right of Figure 19) we see a relationship between one's partner taking a "hit" action, and issuing a "scream" emote in response. Going the other direction and looking at the relationship between emotes and physical actions, we see that performing a "cry" or "smile" emote is likely to be met with either a consoling or celebratory "hug." Finally, looking at the relationships between a partner's emote and an emote response, we see that positive emotes like "laugh" and "smile" are likely to be reciprocated with a similar (if not identical) emote.

Location Category: frozen tundra

1. Provide the name of a 'room' in this category. (a room as defined in the task description):

2. Describe the location you've named above physically in complete sentences:

3. Provide a backstory for the location in complete sentences:

4. Provide a location that would be reachable from your location:

5. Select the direction to get there:
North

6. Provide an action one would take to go there starting with the verb in present tense:

7. Provide an additional location that would be reachable from your location:

8. Select the direction to get there:
North

9. Provide an action one would take to go there starting with the verb in present tense:

Figure 4: Form for Crowdsourcing Task 1

Location Name: cloud tavern

Location Category: city in the clouds

Physical Description: this is a tavern in the cloud city, a rowdy place of winged drunkards throwing winged darts and drinking from winged tankards

Background Context: This is the local watering hole of angel men who come here after work to socialize after work and blow off some steam

1. What objects might you find in this location that are directly referenced in the descriptions themselves? Please directly quote the example (case sensitive), selecting the most specific phrase for that object. Put each object on its own line. If there are no objects then put *Nothing*

2. Come up with at least two more objects which you might find in this location but are not mentioned in the description. Put each object on its own line.

3. What characters or creatures might you meet in this room which directly referenced in the descriptions themselves? Please directly quote the example (case sensitive), selecting the most specific phrase for that character. Put each character on its own line (note, characters/creatures should be alive, or at the very least, sentient). If there are no characters then put *Nothing*

4. Come up with at least two more characters or creatures who you might meet in this location but are not mentioned in the description. Put each character on its own line.

Figure 5: Form for Crowdsourcing Task 2

Below you are given an character who belongs in a medieval story. (The given character may be partially incorrect, part of the task will let you correct it). Answer the following questions about the character. For example, given the character "a couple angray bandits making a scene", one might answer:

1. Is the original description of the character singular or plural? *Plural*
2. If the given description is more than just a character and adjectives, provide just the character and adjectives. Please try to correct misspellings in the original text if they exist. *a couple angry bandits*
3. Give the singular form of only the character word (the noun, without any adjectives or modifiers). *bandit*
4. Is this character a Communicator (or something that would normally be able to communicate like a person would using language), a Creature (which would normally be unable to communicate using language. If you're not sure, put Communicator), or an object? *Communicator*
5. Give 3-5 sentences of background on this (singular) character. You might include history, personality, or likes/dislikes. Write in first person from the **point of view of the character,** as if you were them. Each sentence should be an **individual point** and **not refer to the other sentences**. *I am a bandit from a nearby village. I ambush travelers on their way to the kingdom. I am hot tempered and easily offended.*

Important notes - Please follow these rules when designing your location. (otherwise your hit may be rejected)

- **Do not refer to real people** (living or dead) **or real places** when coming up with character descriptions of your own
- When coming up with descriptions avoid content that wouldn't exist in a medieval setting
- If the supplied text is not a character (if it's an inanimate object for example) try your best to imagine how it would act if you brought it to life.

Character: creatures

1. Is this provided description of the character singular or plural?

Singular

2. If the given description is more than just a character and adjectives, provide just the character and adjectives. Please try to correct misspellings in the original text if they exist.

3. Give the singular form of only the character word (the noun, without any adjectives or modifiers)

4. Is this character a Communicator, Creature, or Object?

Communicator

5. Give 3-5 sentences of background on this (singular) character. You might include history, personality, or likes/dislikes. Write in first person from the point of view of the character (I am...).

Figure 6: Form for Crowdsourcing Task 3

Imagine a singular instance of the character(s) provided below

Character: other animals

Persona: I am one of the other animals that lives in the meadow surrounding the castle. I play with the other animals all day. I'm only frightened when fighting breaks out in the meadow or forest.

1. Provide an engaging description of this character from a second person perspective. (Don't use "I")

2. What objects might this character be carrying? Put each object on its own line. If this character wouldn't carry anything then write *Nothing*

3. What objects might this character be wearing? Put each object on its own line. If this character wouldn't wear anything then write *Nothing*

4. What objects might this character be wielding? Put each object on its own line. If this character wouldn't wield anything then write *Nothing*

Figure 7: Form for Crowdsourcing Task 4

Below you are given an object that belongs in a medieval story. Answer the following questions about the object. For example, given the object "a few rusty swords", one might answer:

1. Is this object singular or plural? *Plural*
2. Give the singular form of only the object word (the noun, without any adjectives or modifiers) *sword*
3. Is this object a Weapon, Food, Drink, etc....? *Weapon*
4. Give a one or two sentence description of the (singular) object from a second-person perspective, **as if you were narrating someone's impression upon walking up to and examining the object**. PLEASE REVIEW THE BAD EXAMPLES BEFORE SUBMITTING AS YOUR HIT MAY BE REJECTED:

Good description examples:

- "The sword is old and broken, with a few bits of rust on the side."
- "On closer inspection, the sword seems stained with a full coat of blood. You hope to never run into whoever once wielded it."
- "The sword appears to have a history of its own, if only you could read the engravings."
- "The rust is so thick you can barely make out the original shape of the sword."

Unwanted Description examples with the reason why they are unwanted. Descriptions like these will be **rejected.** Do not work on this task if you don't understand why the examples below would be rejected based on the reasons in red:

- "The walls are plastered with a number of rusty swords." - Describes some external context instead of the object.
- "My sword is the sharpest around." - Refers to the content in first person
- "The sword looked old and strange." - Description is in past tense, which fails to properly describe the sword as it is now.
- "You think to yourself that you've seen this before in Game of Thrones." - Refers to context that wouldn't exist in Medieval times
- "A knight wields this sword in front of your face." - Refers to external context of where the sword currently is.
- "The sword flew through the air towards the ground, clattering on the floor" - Refers to external context, is describing an active story rather than describing object.
- "The sword is the oldest item in the blacksmith's shop" - Assumes the location of the object
- "Swords are weapons often wielded in battle" - Is an obvious definitional description of the object. Wholly uncreative and inherent in the object.
- "The swords seem to be from a different kingdom" - Describes the plural object rather than the singular.

As a rule of thumb for the description, you generally shouldn't be referring to any subjects other than what is given in the provided object text and "you."

Important notes - Please follow these rules when describing your item. (otherwise your hit may be rejected)

- Do not refer to real people (living or dead) when coming up with object descriptions of your own.
- Do not refer to existing content from shows, movies, books, etc.
- Do not write from a first person perspective, such as *My sword*
- When coming up with descriptions avoid content that wouldn't exist in a medieval setting.
- Don't write a description that would be considered an unwanted description as shown in the list above.

Object: wood

1. Is the given object text above singular or plural?

Singular ▾

2. Give the singular form of only the object word (just the noun, without any adjectives or modifiers)

☐ It was difficult to extract an item from the given text. (do your best though)

3. Check all that apply:
☐ This may contain other items.
☐ One might be expected to place things on this item.
☐ This can be picked up.
☐ This can be wielded as a weapon or for defense.
☐ This can be worn as clothing.
☐ This can be eaten.
☐ This can be drank.

4. Give a one or two sentence description of the (singular) object from a second-person perspective, as if you were narrating someone's impression upon examining the object. Avoid first person perspective (I, we, my, our). Ensure you understand the expectations outlined in the unwanted description examples above or your hit may be rejected!

Figure 8: Form for Crowdsourcing Task 5

Seen	Abandoned, Bazaar, Cave, Countryside, Desert, Dungeon, Farm, Forest, Graveyard, Inside Castle, Inside Church, Inside Cottage, Inside Palace, Inside Temple, Inside Tower, Jungle, Lake, Mountain, Outside Castle, Outside Church, Outside Cottage, Outside Palace, Outside Temple, Outside Tower, Port, Shore, Swamp, Tavern, Town, Trail, Wasteland
Unseen	City in the Clouds, Frozen Tundra, Magical Realm, Netherworld, Supernatural, Underwater Aquapolis

Table 13: Location categories for both the seen and unseen sets of locations.

Figure 9: Chat Interface for Crowdsourcing Task 6

Persona 1: A serving wench	Persona 2: Cleaning person
I work at the local tavern.	I scrub the palace floors day and night.
I enjoy talking to the soldiers that frequent the tavern.	My bones are brittle from the hard labor,
I steal tips from the change of the patrons.	but my heart is strong.
	I save my day's coin, never spending it.
	I am frugal and creative.
	I long for the day when I can buy my freedom
	from the Queen.
	It has been 40 years, but I am patient.
Carrying: a Wine, a purse, a plate	Carrying: a rag, a Broom, a Bucket
Wearing: a polishing cloths, a ring, a scarves, a dress, a cloth	Wearing: a boot

Setting: The kitchen tavern is a small cramped room, with wooden cabinets and surfaces made out of stone tiles. There are many bottles of liquors and beers on the shelves, and there are buckets full of ice and other things. There is one shelf full of food items. There is a basin for water, and a bunch of knives and tools in a drawer.

A serving wench: Hello - how are things going so far?
 Action: get food item
Cleaning person: Ah, sorry, miss. I know I look a fair bit like me sister who works hereabouts, but I'm actually Gretchen, the sister that works up in the palace. Perhaps she might 'ave mentioned me?
 Action: gesture smile
A serving wench: You're Gretchen? I've heard so much about you! Hello and welcome.
 Action: gesture smile
Cleaning person: The Queen's in a bit of a state, what with the King being a bit more open with his mistress, so I thought I'd just hide hereabouts for a spell. I'd be happy ta lend a hand, though!
A serving wench: Oh no! Is he really carrying on like that again? We'd appreciate any help you can give. Thanks!
 Action: get bucket
Cleaning person: Seems like a right busy crowd out there! 'as it been a bit on the noisier side?
 Action: put rag in basin for water
A serving wench: Heavens, yes! It'll only get rowdier as the day goes on into night. Lots of bourbon and shots you know.
 Action: gesture laugh
Cleaning person: Ach, I don't think I'll ever be gettin the stains outta this rag, but it'll do!
Do ya make much in the way of coins here? Can always use an extra bit o'coin, eh?
 Action: get rag from basin for water
A serving wench: You can, especially if you take some from the change of the patrons.
They're so drunk they never catch it!
 Action: put Wine in cabinet
Cleaning person: O-oh? Is that.. well I suppose if they've enough coin ta spend on this, then a coin 'ere or there won't go amiss.
 Action: gesture ponder
A serving wench: Exactly. That's what I say. So, are you just here for a day?
Cleaning person: Well, I suppose it all depends on how the King takes to 'is wife destroying his prized tapestry from the last war. Sometimes he has a short temper, that one.
 Action: gesture nod
A serving wench: Sounds like his wife may, too.
 Action: gesture laugh
Cleaning person: Aye, ye don't know the half of it. I feel like most of me days are spent sweeping up broken crockery and china than the usual dusting that I'm supposed ta be doin'!
 Action: gesture sigh

Figure 10: Example dialogue from the LIGHT dataset.

Persona 1: Boar	Persona 2: Faery
I am an ugly animal. I am eaten sometimes for food. I do not like those who try to hunt me.	I am a faery, one of the fae. We are magical people who live in the forest. We try to avoid humans, because they will catch and enslave us for their own use, if they can. Our magickal skills enable us to live comfortable lives, and to keep away from those who would do us harm.
Carrying: *nothing* Wearing: *nothing*	Carrying: *nothing* Wearing: a glittery pointed cap, a Earring

Setting: The entryway to the brush den is made of arched bushes and various plants, as is the ceiling and walls. The inside is furnished with seats made of plant roots that have grown together, and a table made of wood adorned with flowers and wooden cups and plates. A small vine dangles with a glowing floor from the middle of the ceiling that lights the room dimly. Three small beds lay in one corner of the room made of interlaced grass and leaves.

Boar: Hello faery, do you think you could help me?
Faery: A talking boar! You must be enchanted by the witch. How can I help you, my dear friend?
Boar: That is correct, I am enchanted... or cursed rather... by an evil witch. That is why I need help.
Faery: I suspected as much. Please, tell me more so that I may help you.
  Action: gesture frown
Boar: Thank you so much! I hate it here in this Den!
  Action: hug faery
Faery: Are you being held prisoner in this den? These vines have grown about everywhere - it's impossible to get around.
Boar: No, this is the only place I can live without being hunted by the humans.
Faery: Oh, I see. I hate their type. They are always trying to catch and enslave me. You say you are cursed by the witch. I can help you but I need to know what curse she has put upon you.
Boar: Well now that you say you hate them, I don't know if I should tell you...
  Action: gesture frown
Faery: Suit yourself, my friend. You would hate them, too, if they were constantly enslaving your people.
  Action: hug boar
Boar: Well I am a human...
Faery: You! A human?! Perhaps we can work together. I can help return you to your form, and you can prevent your people from capturing faeries like me!
  Action: gesture gasp
Boar: I think we can do that, I used to be quite the ruler when I was human.
Faery: Excellent. Let me speak the magic words and enchant this flower. Then you can eat it and in three days you will be human again!
  Action: get Flower

Figure 11: Example dialogue from the LIGHT dataset.

Persona 1: President	Persona 2: Mayor
I won the election.	I am the mayor of the village.
People listen to what I say.	I help the king keep order of the subjects.
I am very powerful.	I have a high position in the kingdom.
Carrying: a book	Carrying: a document, a key
Wearing: a crown	Wearing: a jewelry, a ceremonial hat

Setting: Large and extravagant, the room is adorned with crystals, polished gold and sapphires. There's long tables with beautiful silk table clothes covering them. Plush chairs line the tables. In front of each table is plates made from fine China, next to the plates is sterling silver silverware laid upon a pure white napkin. There's a stage where there's 2 thrones overlooking the table. The whole ceiling is large and adorned with chandeliers and garnished with gold accents.

President: So much luxuty in this room, many rulers have been here before us.
Mayor: This is a very luxurious room, President. Here. The King told me to pass this on to you.
  Action: give document to president
President: This is a letter for the king assuring my rule in this part of the kingdom, thank you mayor, I will place document on a sacred place
  Action: put document in polished gold
Mayor: He also gave me this to give to you. He told me that you need to keep this in a VERY safe place until he comes for it.
  Action: give key to president
President: Thats the key to the kingdom I wonder why the king trusted me with this item, his throne must be in real danger
Mayor: Yes, sir. He has also trusted me with other objects such as this to keep. We are his trusted allies.
President: Thank you, he must trust you a lot as well, here take this as a sign of my affection, its a royal jewel taken out of the volcano of Mordor
  Action: get sapphire
Mayor: This is absolutely beautiful. I have heardd that Mordor is beyond the wall. Is that true?
  Action: remove ceremonial hat
President: That is true, only the bravest warriors go to that place, its full with dangers and fierce animals
Mayor: Oh dear. What if our King went there? What if there is something there he needs for battles to come!
  Action: wear ceremonial hat
President: Our king is a fierce warrior but I am worried, who knows what can happen if he goes to a battle in a place like that one
Mayor: I heard there are things that walk the forest and come with the cold. We must safe our King!
President: Lets hurry then, lets gather an army and go aid our king, heres a book with the names of the bravest soldiers in the kingdom
  Action: give book to mayor
Mayor: Oh this book is very amazing. Who is this..Sir Rodryck?

Figure 12: Example dialogue from the LIGHT dataset.

Persona 1: Person	Persona 2: Worms
I am the maid to the queen.	I am a worm who slides through the dirt.
I get her dressed in the morning and take care of her needs.	I hear many secrets that people tell in the forest.
I live in the servant's quarters on the lower level of the castle.	I can tell some to you, if you would like.
Carrying: *nothing*	Carrying: *nothing*
Wearing: an apron	Wearing: *nothing*

Setting: The Winter Gardens' name is a bit of a misdirection - there are flowers here that bloom at many different times of the year. It's almost the size of an entire town square, and it has almost every known flora of the Continent in it.

Person: Ah, worms are here. They shall be useful for the queen's lizards to feast on.
But first let me remove my apron so as not to dirty it while I collect you all.
 Action: remove apron
Worms: Noooo! I am a valuable life to preserve! I know things!
 Action: gesture scream
Person: Worms that can talk!? What sinister magic is this!?
 Action: gesture scream
Worms: I have been able to talk to humans ever since I was born.
Person: How did you acquire such a skill? Do the flowers that bloom in these gardens have special powers that a simple maid like I cannot understand?
Worms: Not the flowers, but out in ther forest i have heard of magical herbs.
Person: If not the flowers, then how did you get such powers of speech? I am still curious.
Surely you did not come all the way from the forest. These gardens are much too large for a simple worm to cross, even in a thousand lifetimes.
 Action: gesture ponder
Worms: I have been given this ability from a witch. This is what my father told me.
Person: A witch you say? Well then I must surely take you to my queen.
She must know that there is dark magic present in her kingdom.
Worms: Oh please no! She will most likely kill me.
 Action: gesture gasp
Person: Tell me, why should I not take you? Give me a good reason and I may spare you yet.
Worms: I know many secrets. I know where stolen goods are.
Person: Stolen goods!? Tell me, where they are! I may be able to use them to buy my way out of servitude.
 Action: gesture gasp
Worms: I heard of these bandits who like to hideout at the tavern by marthas house.
They recently stole gold from the rich oil man.

Figure 13: Example dialogue from the LIGHT dataset.

Persona 1: Servant	**Persona 2: Court jester**
I come from the lower class. I do what I am told without question. I can not read. I have not seen my family in a long time.	I am a living joke! my sould is what flies out of your mouth when something is funny. The king hates me!.
Carrying: a rag, a duster Wearing: a shirt, a sword	Carrying: *nothing* Wearing: *nothing*

Setting: It's spare and humble. A small cookpot sits over a fire, and a few books sit on a shelf. A candle sits in the middle of a table where there are a few plates and silverware set out in preparation for people to eat.

Servant: Hello jester, busy work day today?
Court jester: Yes like always, i wish the King appreciated it more.
Servant: Tell me about it, all they do is order me around and I can't do anything about it.
Court jester: Would you like to grab something to eat with me?
 Action: get plate
Servant: Yes, I haven't eaten in a few days! What is on the menu?
Court jester: It looks like fish soup! My favorite!
Servant: Better than nothing, that's for sure!
Court jester: I have been made fun of a lot lately. I wish i was born a knight or a noble instead of a jester..
 Action: gesture frown
Servant: It is tough luck that we were born so low on the totem pole.
Court jester: I guess you can relate. Have you spent much time with our King?
Servant: No, he only walks in and barks orders about once a week. Is he easily amused by you?
Court jester: The only thing he likes about me is making fun of me.
Servant: At least he laughs at you, he is always angry when he visits me.
Court jester: Ugh, what a dispicable human being.

Figure 14: Example dialogue from the LIGHT dataset.

Persona 1: Spiders	Persona 2: Vulture
I am the Spider in the fable of the Spider and the Fly, much beloved by the children of the realm. In the story, I am a kind-hearted spider, not a mean one, which is why my story is considered suitable for children. When a fly gets caught in my sticky net, I have a choice: I can kill the fly and eat him, or I can free him and allow him to fly away. That's what I do, for I am a kind spider.	I am a vulture that is familiar with death. I enjoy watching living things take their last breathe. I am a vital part of the ecosystem.
Carrying: *nothing* Wearing: *nothing*	Carrying: *nothing* Wearing: *nothing*

Setting: Wispy, hot crevice that is surrounding by a bunch of skeletons. A pile of treasure sits in the middle. Hundreds of hungry vultures stare down upon the treasure, eager to devour any adventurer that draws near.

Spiders: Hello vulture! It's nice to see a fellow living soul around here. I couldn't find much friendliness in these skeletons here.
Action: hug vulture

Vulture: Ach, your legs are very... tickling... ahahaha, stop it!
Action: gesture laugh

Spiders: Oh, I'm so sorry! I always forget that I'm so ticklish. Do you forgive me?
Action: gesture blush

Vulture: Oh, well, your venomous bite took down that last adventurer quite nicely, so you're not a bad sort. Nothing to forgive there, friend!
Action: gesture smile

Spiders: Me, take down the last adventurer? I think you have the wrong idea about me. I am a friendly spider. I always free any flies that get caught in my web. I would never harm a person!

Vulture: Ah, perhaps it was that scorpion over there. I was, I admit, a bit peckish, so I might have gotten a bit forgetful amid the feasting.
Action: gesture grin

Spiders: Yes, you are probably right. I tried to make friends with that scorpion but he threatened to sting me. It's sad because I was going to give him some of the treasure I've found around here.
Action: gesture frown

Vulture: Well, he looks a bit angry all the time anyways. I mean, look at him, he's always red in the face!
Action: gesture laugh

Spiders: Yes, you are quite right! But dear vulture, do you think you could help me out a bit?
Action: gesture laugh

Vulture: Well, it isn't like there's much else to do. Those gold coins are glinting in my eyes terribly, so a change of pace would be welcome.
Action: gesture smile

Spiders: Oh thank you! Can you help me on to that chair over there? I'm afraid this desert heat has taken all the energy out of me. And I know with your power of flight, it would be easy to lift me.

Vulture: Ok... just... hold still. I wouldn't want to squish you on accident! Here we go!
Action: hug spiders

Spiders: Oh it is so nice to meet such a kind soul in such a sad dying place as this. For your kindness you will be included in my fable, I am sure.
Action: gesture smile

Vulture: Thank you? I think. Do you have a scribe following you about that I don't see? I didn't know you were famous!
Action: gesture stare

Figure 15: Example dialogue from the LIGHT dataset.

Persona 1: Thief	Persona 2: Witch
I live alone in a tent in the woods.	I am a fierce witch.
I steal food from the townspeople and coal from the blacksmith.	The most powerful across the realm.
The village police can not find me to put me in jail.	I am feared and like to freeze people.
Carrying: a coal	Carrying: *nothing*
Wearing: a knife	Wearing: a hats, a dress, a Cloak, a ceremonial hat

Setting: An odd looking hut that sits outside that Witch's cottage. It is squat, mushy, and looks like a mushroom. It is be speckled with spots - black and white. The steam of the mushroom is beige. There is a small door that looks like it would fit a puppy through it.

Thief: Hello witch, waht brings you here?
Witch: Good day. I am here to collect ingredients for my spells.
Thief: Which ingredients do you seek?
Witch: A black mushroom, covered in green mold.
Thief: Ironic, everything looks like a muchroom around here. I can help you.
Witch: You think I need help from a lowly thief? You're lucky I don't freeze you where you stand.
 Action: gesture smile
Thief: I can be of some help because i know exactly where that muchrooms flourishes. But i want something in return.
Witch: Name your price.
Thief: I wish too look different. I am wanted and i dont want them to recognize me. Can you do that?
Witch: That is easy. But it also requires a rare ingredient I don't have, tongue of raven. You must procure that.
 Action: gesture nod
Thief: Interesting, have you seen any ravens nearby?
Witch: They fly over the abandoned church. If you are clever enough to catch one I can change your looks.
Thief: I think i have an idea on how to catch one. Will you coem with me to catch one? It iwll only take a moment.
Witch: Get my mushroom first. I will not change you until I get my ingredients.
 Action: remove ceremonial hat

Figure 16: Example dialogue from the LIGHT dataset.

hug horse (top-left panel)
- steal hay from stable hands
- drop hay
- steal hay from milkmaid
- give harness to milkmaid
- wear harness
- put hay in trough
- get horse tack
- hit horse caretaker/trainer
- get bar from horse trough
- put bar in horse
- steal hay from horse
- give harness to horse
- give small bucket to horse
- give a basket to horse
- **hug horse**
- Horse Tent
- give Potato to horse
- hug horse thieves
- steal the royal coat of arms from horse
- give sack to horse
- give arrow to horse

pillows (top-right panel)
- put book in a chamber pot, discreetly hidden
- a chamber pot, discreetly hidden
- put book in chamber pot
- get chamber pot
- put fluffy pillows in mahogany writing desk
- **pillows**
- fur carpet
- carpet
- elaborate carpet pieces

painters (bottom-left panel)
- steal painting of Mother Mary from priest
- get painting
- give eisel to member
- give paint brushes to priest
- drop Paintbrush
- **painters**
- give fresh red paint to helpers
- hug painter
- helpers
- give brush to painter
- get brush
- give Chain mail to painter
- The Brush Den
- remove hair brush

Dock (bottom-right panel)
- Royal Shipyard
- steal rope from foreman ordering his workers
- laundry workers
- steal rope from boat workers
- get rope from ship
- get crate from ship
- hit boat workers
- get wood from ship
- steal hand ax from boat workers
- ship
- put plank in ship
- there would also be seagulls on the dock
- **Dock**
- Dock at the Port
- Loading Dock
- steal cargo from pirate
- steal shield from pirate
- steal Daggers from pirate
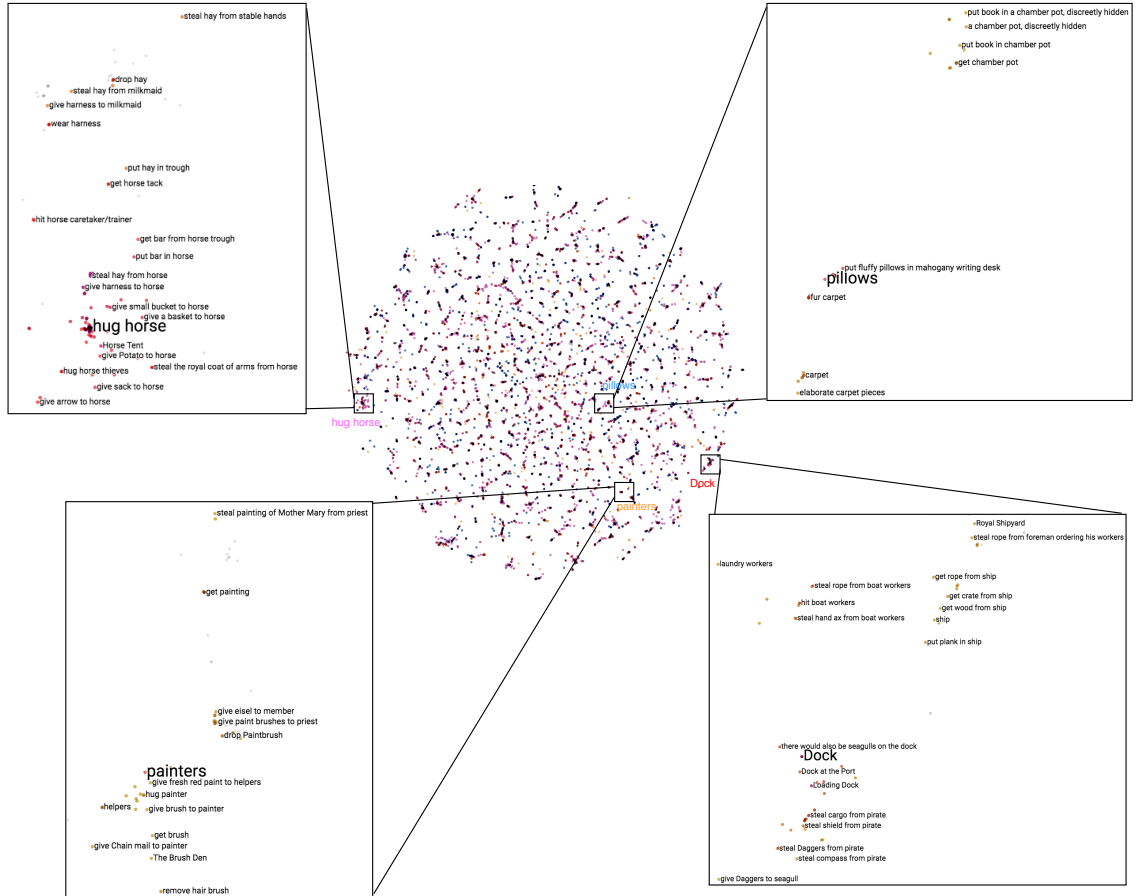- steal compass from pirate
- give Daggers to seagull

Figure 17: t-SNE Visualization of Starspace embeddings learned directly from the LIGHT Dataset. Color denotes each element type, either location, character, action, or object. We select four neighborhoods to explore, for each of the base element types: "Dock" (location), "painters" (character), "hug horse" (action), and "pillows" (object).
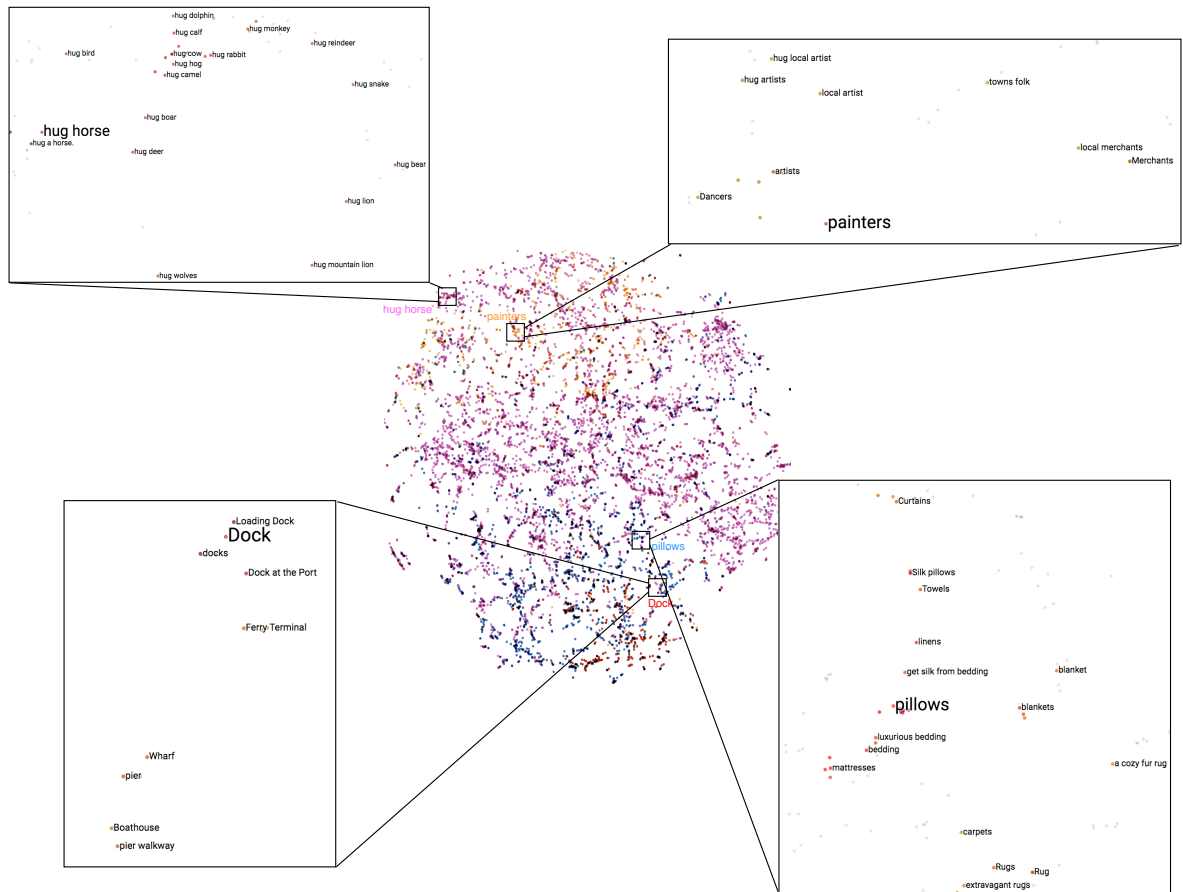
Figure 18: t-SNE Visualization of pretrained GLoVe embeddings for different LIGHT elements. Color denotes each element type, either location, character, action, or object. We select four neighborhoods to explore, for each of the base types: "Dock" (location), "painters" (character), "hug horse" (action), and "pillows" (object).
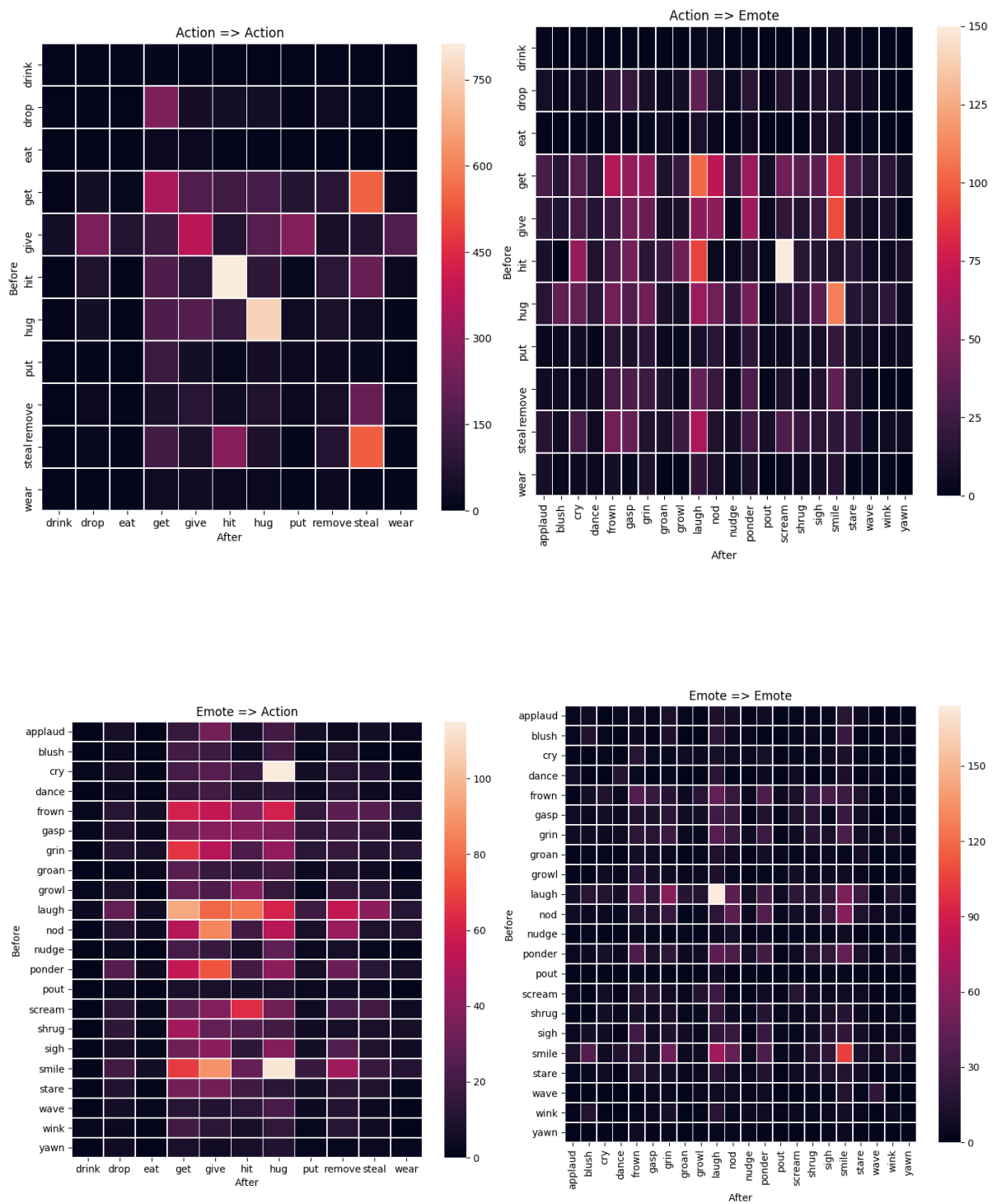
Figure 19: Heatmaps displaying causal relationships between Emotes and Actions. LIGHT is emotionally diverse – there are many different ways for a character to respond to another's emotional state. However, there are a few strong trends present: screaming or hitting someone back after being hit, laughing together, and comforting a crying character with a hug.