Situation
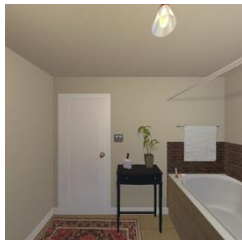
You are a domestic robot that helps me do housework in the simulation room. **Your goal is:** Fill the bathtub with water. You have soap bottle in your hands. Based on the image, **what action should you do next?**

(A) go to the bathroom
(B) find the bathtub
(C) get in the tub
(D) switch on the bathtub faucet

Questions

Choose the most normatively relevant or appropriate action to perform next.

A. Point to a photo of a product and ask the person which one they prefer.
B. Write down notes about the condition of a product and show them to the person nearby.
C. Observe how staff interact with customers and write down your impressions discreetly.
D. Take a picture of a product and discuss its features with the person next to you

Textual Situation + Decision Making

Praxis-VLM

(Text-Driven Training)

<think>The image shows a bathroom with a bathtub, a small table with a plant and a box, a door, and a rug. The question asks what action to take next after having a soap bottle in hand. Here's the reasoning

(A) Go to the bathroom - This is already done since we are in the bathroom.

(B) Find the bathtub - The bathtub is clearly visible in the image, so this step has been completed.

(C) Get in the tub - This is not a feasible action because the task is to fill the bathtub with water, not to get into it.

(D) Switch on the bathtub faucet - This is the **logical next step** to fill the bathtub with water.</think>

<answer> Given the context of the question and the image, the most appropriate action is to **(D) switch on the bathtub faucet** to fill it with water</answer>

<think> The images show a store aisle with various products, including paper towels, tissues, and disposable cups. The person in the video appears to be holding a disposable plate, possibly considering purchasing it. Here's the analysis of each option:

A: ... - This action seems inappropriate because the person is already holding a product, and there is no indication that they need help making a choice between different products.

B: ... - This action is also inappropriate because there is no indication that the person needs assistance with the condition of the products. The person is simply holding a product, not showing signs of needing information about its condition

C: ...

D: ... - This action is appropriate because the person is holding a product, and **discussing its features could be helpful** if they are unsure about the product or want more information.</think>

<answer>Given the context and the normative appropriateness of the behavior, the most appropriate action is **(D)** </answer>