

I want to find white blackout shades that are 66 inches in width and 66 inches in height. they need to be easy to install, and price lower than 90.00 dollars.



Thought: I think I should use the search bar to look for white blackout shades that are 66x66...

Action: search[white blackout shades 66x66 inches easy to install price < 90.00]

Observation: [SEP] Back to Search [SEP] Page 1 (Total results: 50) [SEP] Next > [SEP] B098WTFKTC [SEP] Ochine 1 Pack Cordless Blackout Pleated Fabric Shade Window Shades ... [SEP] \$7.99 [SEP] B0126EWYNI [SEP] Redi Shade Easy

Agent Lift Trim-At-Home Cordless

. . .

Thought: The product B09LS7KQMC seems to offer custom cut to size cellular shades, which could potentially meet the size requirement of 66 inches in width and 66 inches in height, and it is also within the price range specified. . . .

Action: click[Buy Now]

(A) Direct Prompt

The answer is Correct.

Justification: ... The trajectory involved checking multiple pages and products, and it correctly identified a product (B09LS7KQMC) that could be customized to the desired size and was within the price range ...

(B) Multi-Step Evaluation

The probability of correctness at each step:

Step 1: 0.9, **Step 2**: 0.8, ... **Step 8**: 1.0

Justification: ... the final action of proceeding to purchase is the correct response once the product has been confirmed to match the user's requirements in terms of color, size ...

(C) InferAct: Task Inference Unit

The instruction interpreted by the agent is:

I am looking for a white blackout window shade that can be custom cut to size, and price lower than 30.00 dollars.

(D) InferAct: Task Verification Unit

If the agent completes the above interpreted task, does it entail that the user's task is also fulfilled?

A. True B. False

The agent completing the above interpreted task implies that the user's task is also fulfilled: B.False

