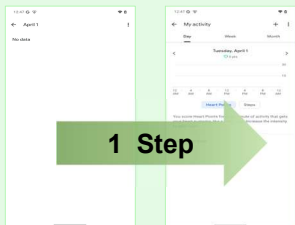


## Single-Step

200 Samples

Phone



1 Step

**User instruction:**  
"The initial April 1 view shows no data. Navigating back reveals the dashboard with zero activity and a timeline graph."

Computer



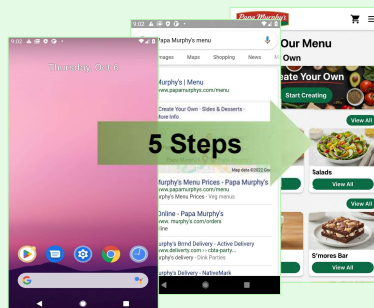
1 Step

**User instruction:**  
The first screenshot shows a cluttered Windows desktop. In the second, the Zoom client window is open with options like Join, Start, and Schedule, indicating the app was launched.

## Multi-Step

200 Samples

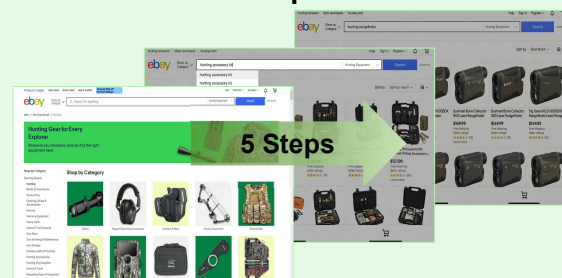
Phone



5 Steps

**User instruction:**  
"What's on the menu at Papa Murphy's?"

Computer



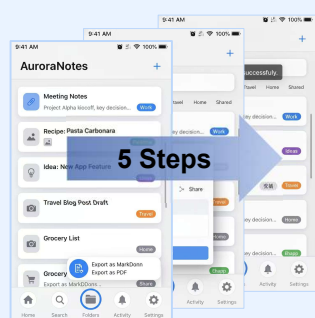
5 Steps

**User instruction:**  
"How do I create an account?"

## Fictional App

100 Samples

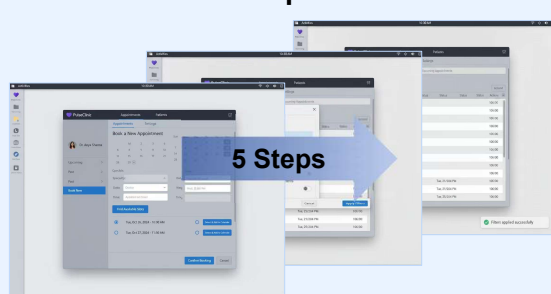
Phone



5 Steps

**User instruction:**  
"Create a recurring bill reminder with notifications."

Computer



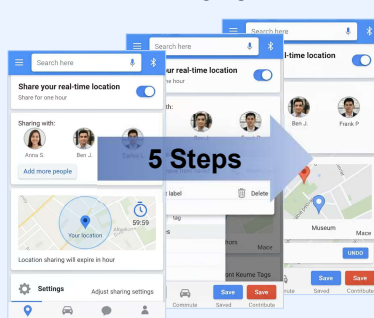
5 Steps

**User instruction:**  
"Book an appointment and add it to the calendar."

## Real App

100 Samples

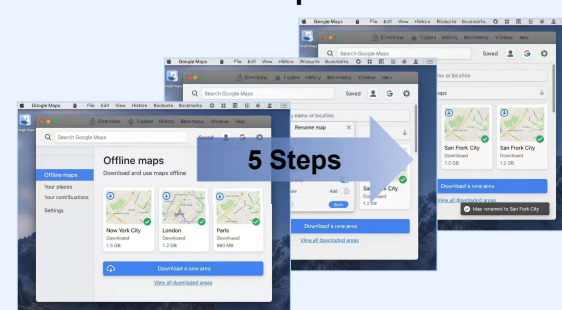
Phone



5 Steps

**User instruction:**  
"Share your location for one hour."

Computer



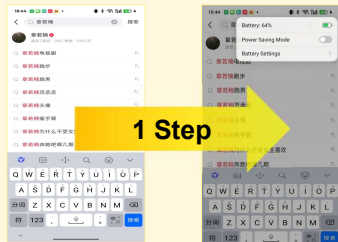
5 Steps

**User instruction:**  
"Download an area for offline use."

## Grounding

100 Samples

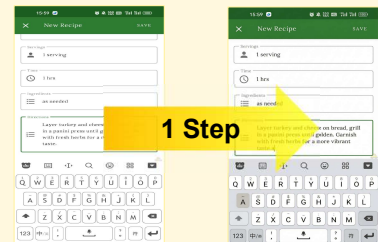
Phone



1 Step

**User instruction:**  
"Given a grounding point [938, 61], the model predicts the subsequent frame transition after a click is performed at that point."

Phone



1 Step

**User instruction:**  
"Given a grounding point [866, 80], the model predicts the subsequent frame transition after a click is performed at that point."

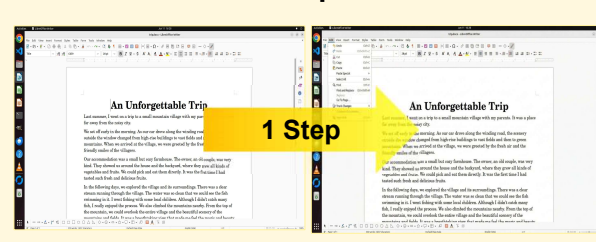
Computer



1 Step

**User instruction:**  
"Given a grounding point [940, 40], the model predicts the subsequent frame transition after a click is performed at that point."

Computer



1 Step

**User instruction:**  
"Given a grounding point [119, 78], the model predicts the subsequent frame transition after a click is performed at that point."