Although these strategies are *latent cognitive constructs*, they can be **inferred from observable gaze behavior**, especially with theoretical guidance from:

* **Lohman’s distinction** between *piecemeal* and *heuristic/global* strategies (strategy function)
* **Hegarty’s empirical indicators** of strategy in scanpaths and fixation patterns (strategy execution)

**✅ Step-by-Step Guide to Annotating Strategies**

**🔹 Step 1: Define Strategy Labels Based on Theory**

| **Strategy** | **Lohman / Hegarty Basis** | **Core Behavioral Indicators** |
| --- | --- | --- |
| **Planner / Heuristic** | Goal-directed, global processing | Low entropy, short scanpaths, few revisits, early target fixation |
| **Explorer / Piecemeal** | Trial-and-error, local processing | High entropy, many AOI switches, longer paths, scattered fixations |
| **Checker** | Verification and re-evaluation | Multiple regressions, long dwell time on final AOIs, late re-visits |

These labels can also align with clusters or expert-novice distinctions.

**🔹 Step 2: Select Annotation Metrics**

Use **quantitative gaze metrics** as proxies for cognitive strategy. Examples:

* **Entropy of Gaze Transitions**
* **Fixation Count and Duration**
* **AOI Revisits**
* **Scanpath Length / Regularity**
* **Time to First Fixation on Target**
* **Convex Hull / Gaze Dispersion**

**🔹 Step 3: Create Labeling Heuristics**

Develop **threshold-based rules** or **decision criteria** using pilot data or theory-informed cutoffs.

Example (simplified pseudocode):

python

CopyEdit

if entropy < 1.0 and revisits < 3:

label = "Planner"

elif entropy > 1.4 and scanpath\_length > threshold:

label = "Explorer"

elif revisits > 5 and late\_dwell > threshold:

label = "Checker"

Refine with histograms or scatter plots of actual participant data.

**🔹 Step 4: Apply and Validate**

* Annotate your 150-sample dataset using the heuristics.
* Perform **inter-rater reliability** if manual review is involved.
* Optionally use **clustering (unsupervised)** to validate or discover additional subtypes.

**🔁 Bonus: Hybrid Labeling Pipeline**

You can combine theory + ML like this:

1. **Manually label a subset** using heuristics.
2. **Train a classifier (e.g., Random Forest)** on those labels.
3. **Automatically label the rest** of the data.
4. **Interpret feature importance** to validate alignment with theory.

A strong starting point for designing your strategy-annotation plan is **Hegarty, Stieff & Dixon (2015)**:

*“Reasoning with Diagrams: Towards a Broad Ontology of Spatial Thinking Strategies.”*  
This paper outlines key spatial-thinking strategies and provides theoretical justification, including **strategy types and processes**, making it ideal for grounding annotation criteria [researchgate.net+11hegarty-lab.psych.ucsb.edu+11researchgate.net+11](https://hegarty-lab.psych.ucsb.edu/publications?utm_source=chatgpt.com).

Also consider **Nazareth, Odean & Pruden (2016)**:

*“The Use of Eye‑Tracking in Spatial Thinking Research.”*  
This chapter thoroughly reviews how eye tracking reveals **strategy selection and flexibility**—including piecemeal versus heuristic processing patterns—and discusses associated gaze metrics like fixations, transitions, and entropy. It’s explicitly framed to connect eye-tracking data with theoretical strategies [pubs.acs.org+7researchgate.net+7researchgate.net+7](https://www.researchgate.net/publication/308995766_The_Use_of_Eye-Tracking_in_Spatial_Thinking_Research?utm_source=chatgpt.com)[files.eric.ed.gov+4scribd.com+4pubmed.ncbi.nlm.nih.gov+4](https://www.scribd.com/document/793259636/The-Use-of-Eye-Tracking-in-Spatial-Thinking-Research?utm_source=chatgpt.com).

**📋 Why These Two Make Ideal Starting Points**

| **Paper** | **Why It's Ideal for Annotation** |
| --- | --- |
| Hegarty et al. (2015) | Establishes a **taxonomy** of spatial strategies; theory-driven with clear categories. |
| Nazareth et al. (2016) | Maps spatial strategy constructs to **concrete eye-tracking metrics**, enabling operational labels. |

**🧠 How to Use These Papers for Annotation**

1. **Define Strategy Categories**  
   Use Hegarty et al.'s taxonomy to identify which strategies your lab tasks engage (e.g., *visual comparison*, *mental rotation*, *piecemal processing*).
2. **Extract Gaze Features**  
   From Nazareth et al., choose gaze metrics that correspond closely to strategy behaviors—e.g., transition entropy for exploration, dwell time for depth, first-fixation latency for prioritization.
3. **Formulate Labeling Heuristics**  
   Based on these mappings, set threshold-based rules or visuals that assign strategy labels grounded in theory:
   * *Low entropy + early target fixations = heuristic/planner*
   * *High entropy + scattered fixations = piecemeal/explorer*
4. **Validate and Iterate**  
   Annotate a subset of data, then use expert review or inter-rater reliability to refine definitions and thresholds.

These two works provide the **theoretical and methodological scaffolding** you need to annotate gaze data in a way that is:

* **Grounded in established educational research**
* **Supported by operational gaze metrics**
* **Appropriate for supervised machine learning research**

**📚 1. Hegarty, Stieff & Dixon (2015) – “Reasoning with Diagrams: Towards a Broad Ontology of Spatial Thinking Strategies”**

* **Content:** Presents a taxonomy of spatial reasoning strategies such as *mental rotation*, *visual comparison*, and *logical inference*—differentiating between piecemeal (analytic) and holistic (heuristic) approaches [researchgate.net+15hegarty-lab.psych.ucsb.edu+15researchgate.net+15](https://hegarty-lab.psych.ucsb.edu/publications?utm_source=chatgpt.com).
* **Use for annotation:** Provides clear theoretical categories to map onto gaze behavior (e.g., analytic strategies ↔ focused AOI fixations; holistic strategies ↔ global scanpaths).

**📘 2. Nazareth, Odean & Pruden (2016) – “The Use of Eye‑Tracking in Spatial Thinking Research”**

* **Content:** Reviews how eye-tracking reveals strategy selection and flexibility in spatial tasks, linking metrics like fixation patterns, transition entropy, and dwell time to cognitive strategies [scribd.com](https://www.scribd.com/document/793259636/The-Use-of-Eye-Tracking-in-Spatial-Thinking-Research?utm_source=chatgpt.com).
* **Use for annotation:** Offers concrete guidance on which gaze metrics correspond to piecemeal versus heuristic processing—essential for building your labeling rubric.

**🧭 How to Use These Papers for Annotation Planning**

| **Step** | **Action** |
| --- | --- |
| **1. Define cognitive strategies** | Use Hegarty et al. (2015) to define labels like *analytic/piecemeal* vs *holistic/heuristic*. |
| **2. Map gaze features to strategy** | Draw from Nazareth et al. (2016) to identify metrics: entropy, first-fixation latency, dwell time, AOI visits. |
| **3. Formulate labeling heuristics** | E.g., |

* **Piecemeal**: high transition entropy, many short fixations
* **Heuristic**: low entropy, early target fixation, regular scanpaths |  
  | **4. Pilot & iterate** | Annotate a small sample set, refine thresholds or decision rules guided by gaze distributions and expert review. |

**🧩 Mapping Table: Gaze Metrics to Cognitive Strategies**

| **Strategy Label** | **Cognitive Basis** | **Expected Gaze Pattern** | **Indicative Eye-Tracking Metrics** | **Heuristic Rule (Example)** |
| --- | --- | --- | --- | --- |
| **Heuristic / Planner** | Global processing, high skill (Lohman, Hegarty) | Structured scanpaths, early key AOI attention | - Low entropy - Few AOI transitions - Short scanpath - Early fixation on key AOIs | if entropy < 1.0 and TFF < 2000 ms: → Planner |
| **Piecemeal / Explorer** | Local/element-wise reasoning (Lohman) | Scattered fixations, many AOI switches | - High transition entropy - High fixation count - High AOI revisit count - Long scanpath | if entropy > 1.4 and AOI\_revisits > 5: → Explorer |
| **Checker** | Verification, uncertainty (Hegarty) | Frequent re-scanning of AOIs, late-stage fixation | - Many AOI revisits - High dwell time on AOIs - Late fixations on targets | if AOI\_revisits > 5 and dwell\_time > threshold: → Checker |

These heuristics are customizable based on your actual data distribution (you can adjust entropy cutoffs or fixation counts based on histograms or z-scores).

**🧪 Code Template (Python / Pandas + Rule-Based Annotation)**

python

CopyEdit

import pandas as pd

def annotate\_strategy(row):

if row['entropy'] < 1.0 and row['time\_to\_first\_fix'] < 2000:

return "Planner"

elif row['entropy'] > 1.4 and row['AOI\_revisits'] > 5:

return "Explorer"

elif row['AOI\_revisits'] > 5 and row['dwell\_time'] > 3000:

return "Checker"

else:

return "Unclassified"

# Load your preprocessed gaze data (150 samples)

df = pd.read\_csv("eye\_tracking\_features.csv")

# Apply annotation

df['strategy\_label'] = df.apply(annotate\_strategy, axis=1)

# Save or review annotated data

df.to\_csv("annotated\_strategies.csv", index=False)