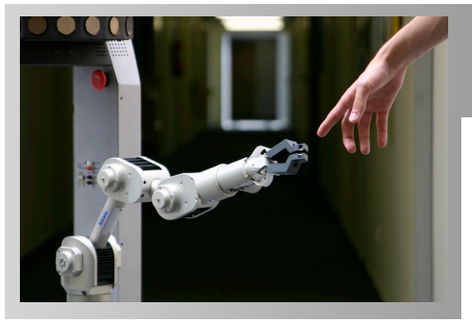


Situated Interaction Testbed

Geert-Jan M. Kruijff
Cognitive Systems @ DFKI LT
gj@dfki.de

Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Purpose of the testbed

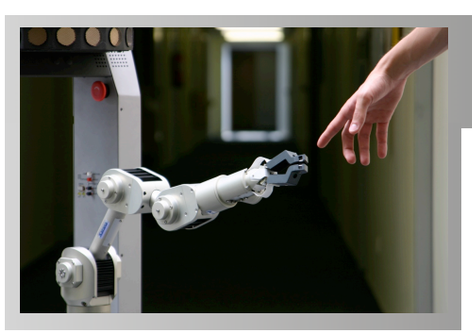
- A collection of visual scenes
  - Dimensions for systematic variation of scenes
  - Base level dimensions defined at level of individual objects
- Use of visual scenes
  - Basis for verification of hypotheses about factors influencing scene comprehension
  - Basis for investigation what are primitive features in which scenes are (can be) characterized

Object templates

- Function of object templates
  - An object template defines absolute as well as context-sensitive characteristics that identify the object
  - For each characteristic, potential distractors in the scene are identified
- Form of object templates
  - Absolute physical characteristics:
    - properties: color, shape
    - type
  - Scalar material characteristics
    - physical properties: size
  - Context characteristics:
    - spatial relations: ego-topological, ego-projective, allo-object-topological, allo-object-projective, allo-speaker-topological, allo-speaker-projective
    - salience (ego-centered): visual salience, topokinetic salience

Scene templates

- Matrix organization of a scene
  - Spatial resolution at 10cm x 10cm cells
  - Matrix grid organization over cells: e.g. 5 by 4 cells, to yield a 50cm x 40 cm workspace
- Spatial organization of scenes
  - One object per cell
  - Discrete approximation of topological, projective spatial relations
  - Saliency measures can be directly defined on matrix grid cells

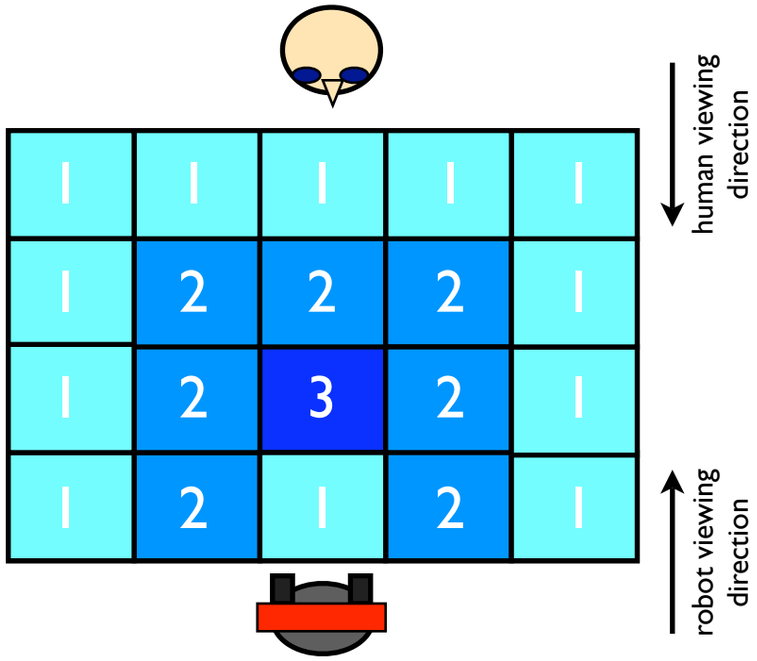


Object Manipulation Workspace

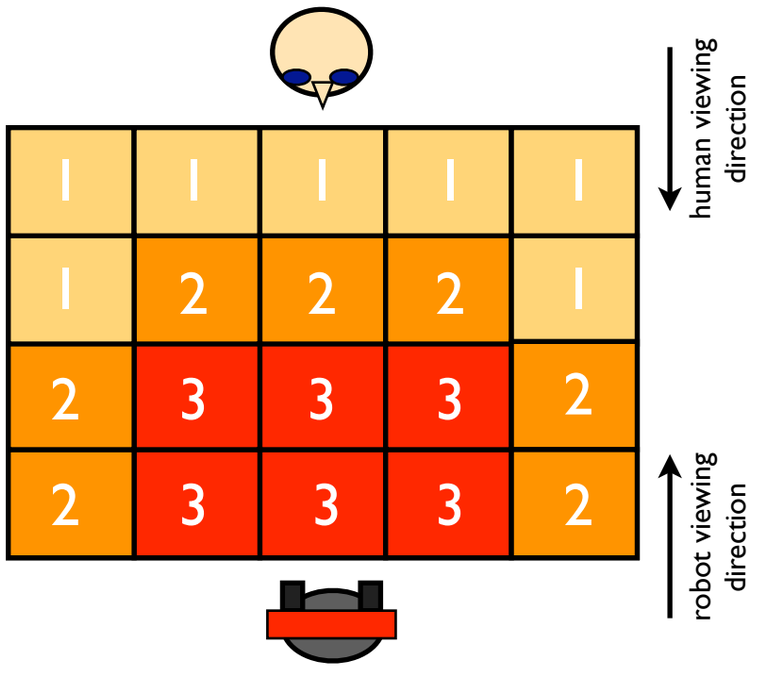
Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Scene, object definitions

scene



visual salience scoring



topokinetic ("reachability") salience scoring

object

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Scene description

1	1	1	1	1
1	2	2	2	1
1	2	3	2	1
1	2	1	2	1

1	1	1	1	1
1	2	2	2	1
2	3	3	3	2
2	3	3	3	2

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

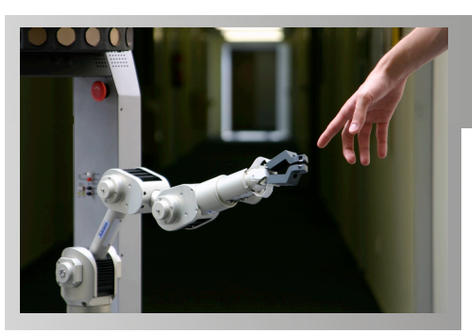
Characteristic	Value	Distractors
Type		
Color		
Shape		
Size		
Topological		
ego-centered		
object-centered		
human-centered		
Projective		
ego-centered		
object-centered		
human-centered		
Visual salience		
Kinetic salience		

Systematic variation

- Scene variation by salience and physical properties
  - Number of potential distractors by type, color, contrastive scalar attribute
  - Number of potential distractors by visual salience and {type, color, scalar}
  - Number of potential distractors by kinetic salience and {type, color, scalar}
  - This enables variation in the identification of objects by individual properties
- Scene variation by spatial relations
  - In addition to potential distractors arising from individual properties, we can also vary the number of distractors relative to spatial descriptions
  - Number of potential distractors by topological relationships
  - Number of potential distractors by projective relationships
  - This enables variation in identification of objects thru descriptions with a higher cognitive load (attributes < topological < projective)

Templates for variation in physical properties

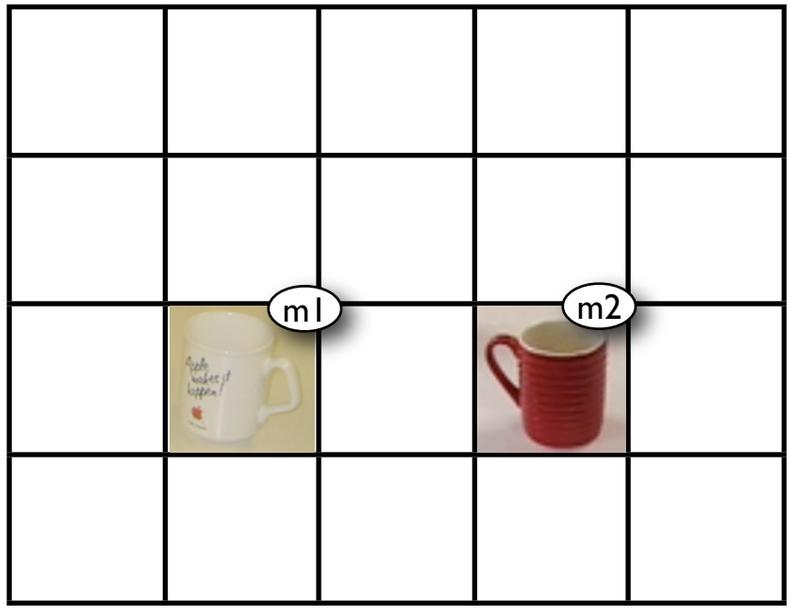
1. Identification by distinctive material property
   - Distractors with same type but variation in material properties (shape, color)
2. Identification by contrastive scalar property
   - Distractors with same type but variation in scalar property (absolute: big/small; comparative: bigger/smaller - which requires an absolute “big” to be definable too)
3. Identification by visual salience and type
   - Distractors with same type but varying salience
4. Identification by visual salience, type, and distinctive material property
   - Distractors with same type and salience, but varying material properties (color, shape)
5. Identification by visual salience, type, and contrastive material property
   - Distractors with same type and salience, but varying scalar property (absolute or comparative)
6. Identification by kinetic salience and type
   - Distractors with same type, and graspable, but varying kinetic salience
   - Distractors with same type and identical kinetic salience, but varying in graspability (yes/no)
7. Identification by type, and combined salience
   - Distractor sets by salience type each have more than 1 distractor, but intersection between the distractors is empty (i.e. use both salience characteristics to resolve the referent)




Identification by distinctive material property

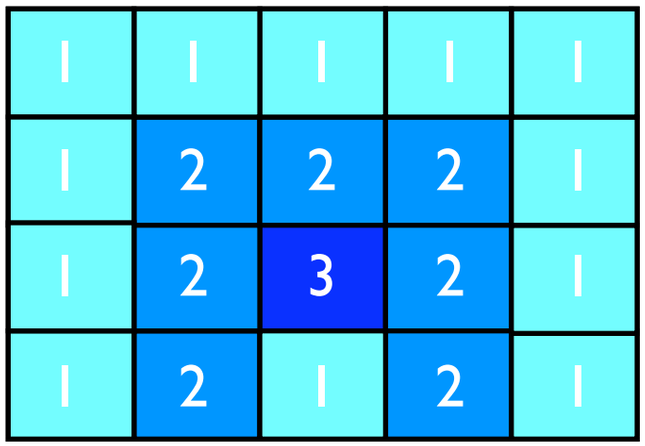
Scenes MP1-4

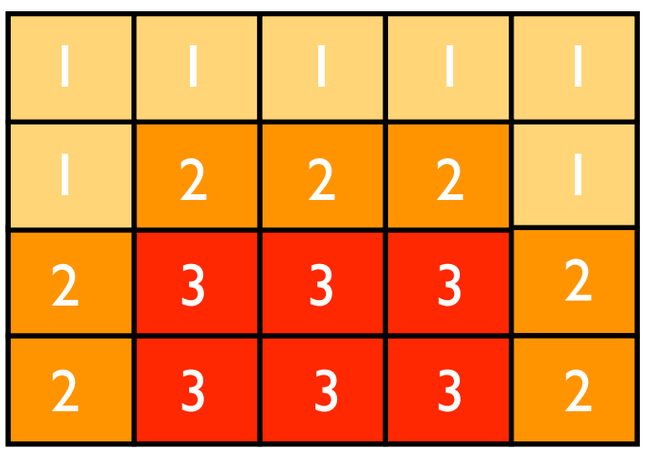
Scene#MP.1



m1		
Characteristic	Value	Distractors
Type	mug	m2
Color	white	∅
Shape	container	m2
Size	--	--
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	3	m2

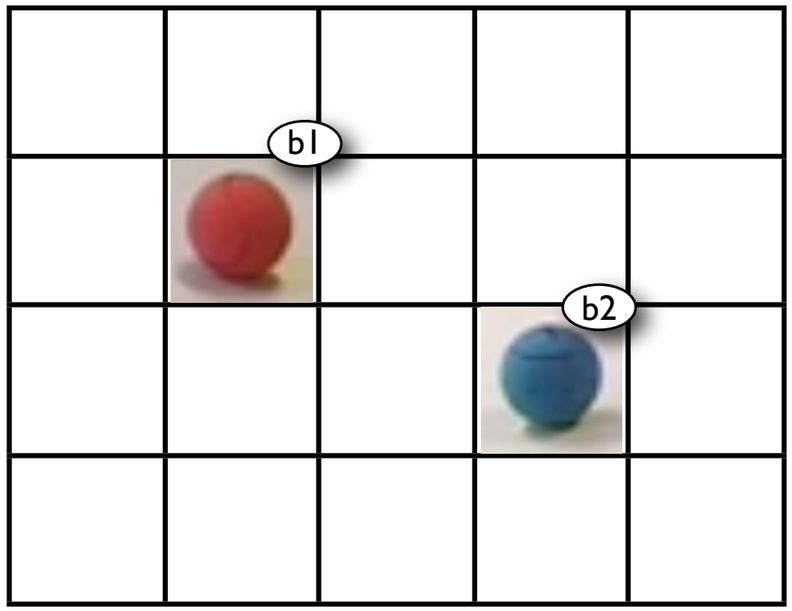
m2		
Characteristic	Value	Distractors
Type	mug	m1
Color	red	∅
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	m2





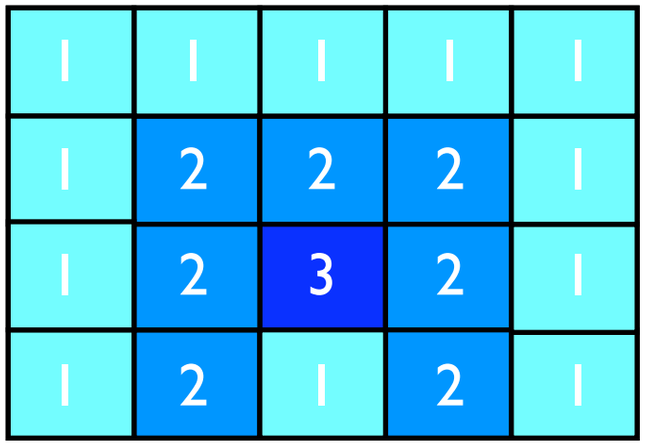
- Identification by distinctive material property
  - Distractors with same type but variation in material properties
  - The individuals can be identified directly through distinct colors
  - E.g. “the white mug”, “the red mug”
  - (Additional distinctive properties: projective spatial relations)

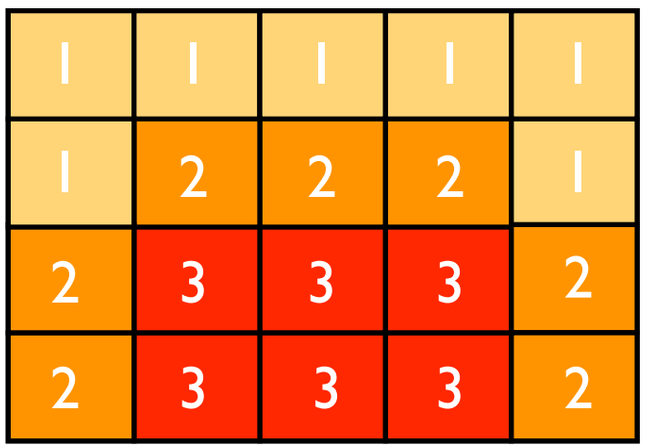
Scene#MP.2 - data



b1		
Characteristic	Value	Distractors
Type	ball	b2
Color	red	∅
Shape	sphere	b2
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	b2
Kinetic salience	2	∅

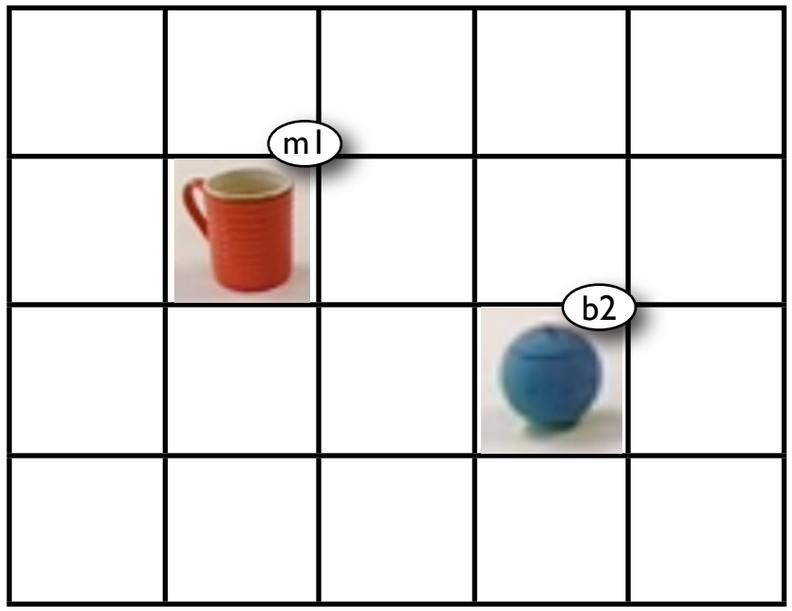
b2		
Characteristic	Value	Distractors
Type	ball	b1
Color	blue	∅
Shape	sphere	b1
Size	--	--
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	b1
Kinetic salience	3	∅





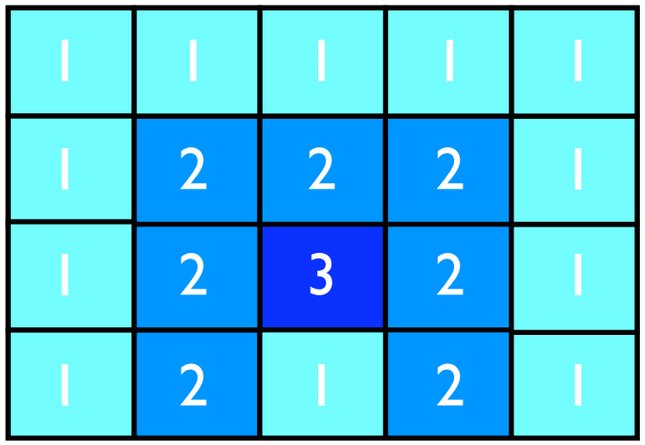
- Identification by distinctive material property
  - Distractors with same type but variation in material properties
  - The individuals can be identified directly through distinct colors
  - E.g. “the white mug”, “the red mug”
  - (Additional distinctive properties: projective spatial relations, kinetic salience)

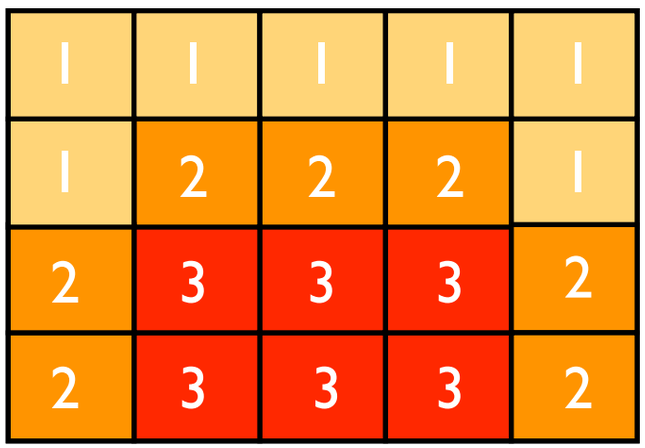
Scene#MP.3 - data



b1 Characteristic	Value	Distractors
Type	mug	∅
Color	orange	∅
Shape	container	∅
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	b2
Kinetic salience	2	∅

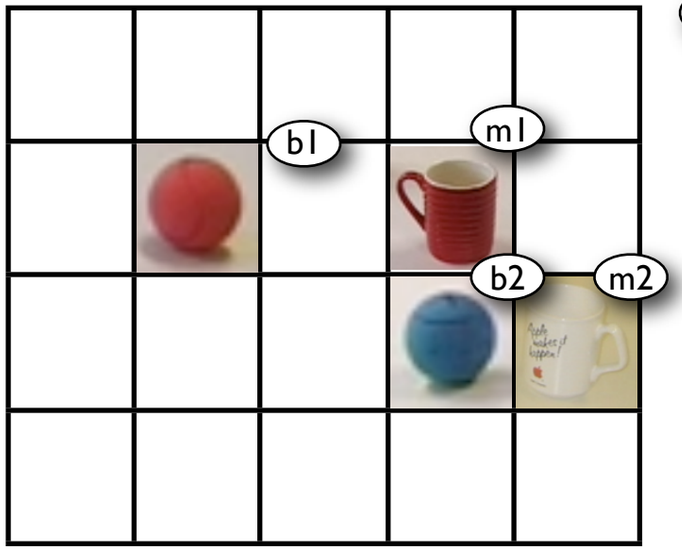
b2 Characteristic	Value	Distractors
Type	ball	∅
Color	blue	∅
Shape	sphere	∅
Size	--	--
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	∅

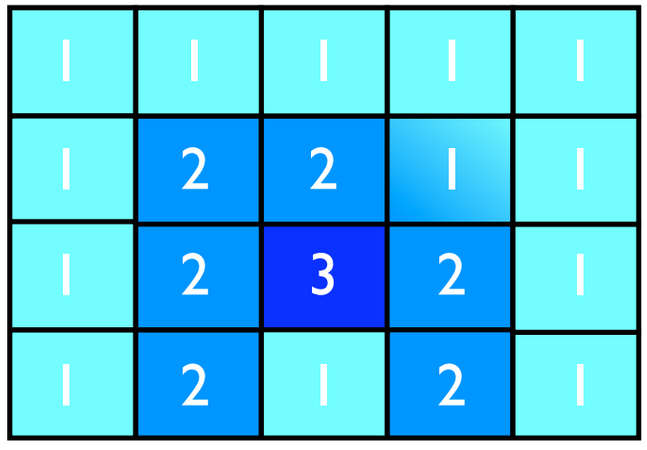


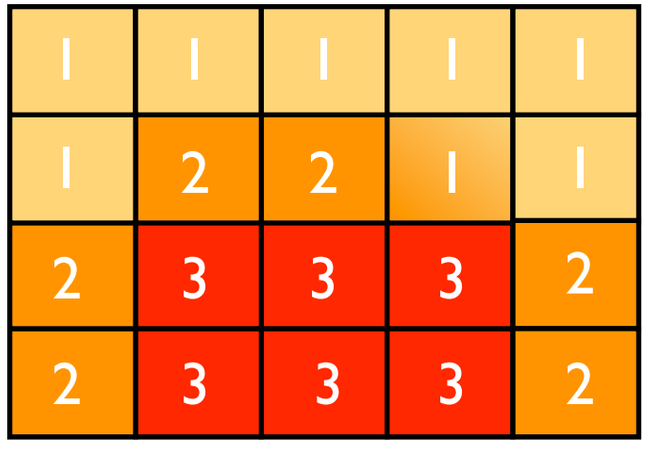


- Identification by distinctive material property
  - The individuals can be identified directly through type
  - E.g. “the mug”, “the ball”
  - (Additional distinctive properties: projective spatial relations, kinetic salience)

Scene#MP.4 - data







b1

Characteristic	Value	Distractors
Type	ball	b2
Color	red	m1
Shape	sphere	b2
Size	--	--
Topological		
ego	far	m1
allo-object	--	--
allo-human	near	m1
Projective		
ego	left	∅
allo-object	--	--
allo-human	right	∅
Vis. salience	2	b2,m1
Kin. salience	2	m2

b2

Characteristic	Value	Distractors
Type	ball	b1
Color	blue	∅
Shape	sphere	b1
Size	--	--
Topological		
ego	near	∅
allo-object	--	--
allo-human	far	m2
Projective		
ego	right	m1,m2
allo-object	--	--
allo-human	left	m1,m2
Vis.salience	2	b1,m1
Kin. salience	3	∅

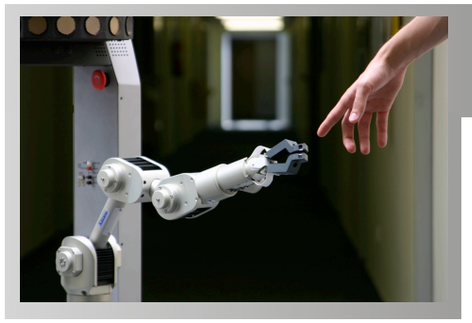
m1

Characteristic	Value	Distractors
Type	mug	m2
Color	red	b1
Shape	contain	m2
Size	--	--
Topological		
ego	far	b1
allo-object	--	--
allo-human	near	b1
Projective		
ego	right	b2,m2
allo-object	--	--
allo-human	left	b2,m2
Vis.salience	1	m2
Kin. salience	1	∅

m2

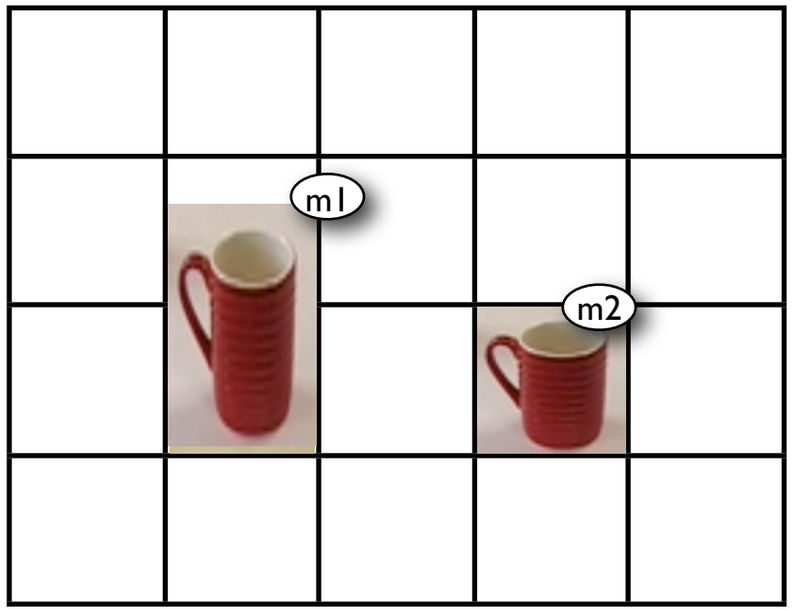
Characteristic	Value	Distractors
Type	ball	b1
Color	blue	∅
Shape	contain	b1
Size	--	--
Topological		
ego	near	∅
allo-object	--	--
allo-human	far	b2
Projective		
ego	right	m1,b2
allo-object	--	--
allo-human	left	m1,b2
Vis.salience	1	m1
Kin. salience	2	b1

- Identification by distinctive material property
  - Distractors with same type but variation in material properties
  - The individuals can be identified directly through a combination of type, and color or shape. E.g. “the white/red mug”, “the blue sphere”
  - Individuals cannot be distinguished by type (“the mug”) or by color (“the red thing”) - there are always distractors
  - (Additionally, non-object centered spatial relations are usually ambiguous, as are saliency measures)
  - (Interestingly, because of occlusion the visual and kinetic saliency of m1 should probably be decreased to 1; “grab the mug” would more likely refer to m2 in this case as it is easier to reach)



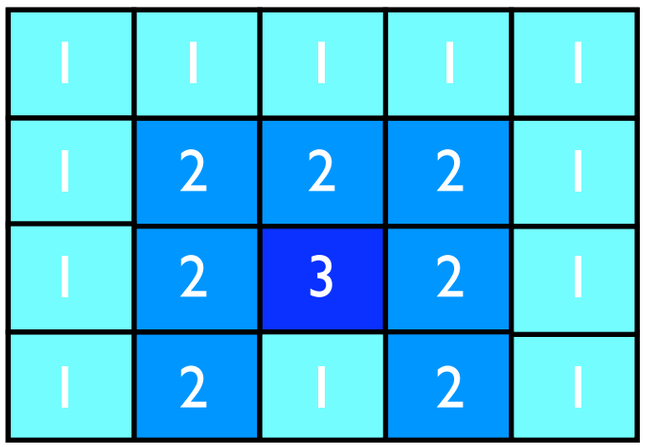

Identification by contrastive scalar property

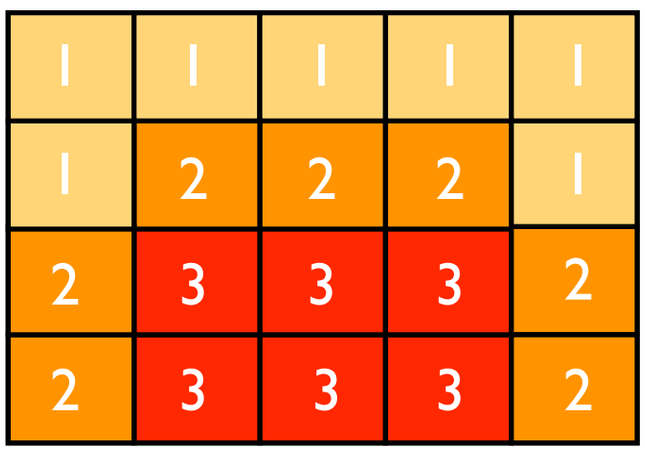
Scene#CP.1



m1 Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	big	∅
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	3	m2

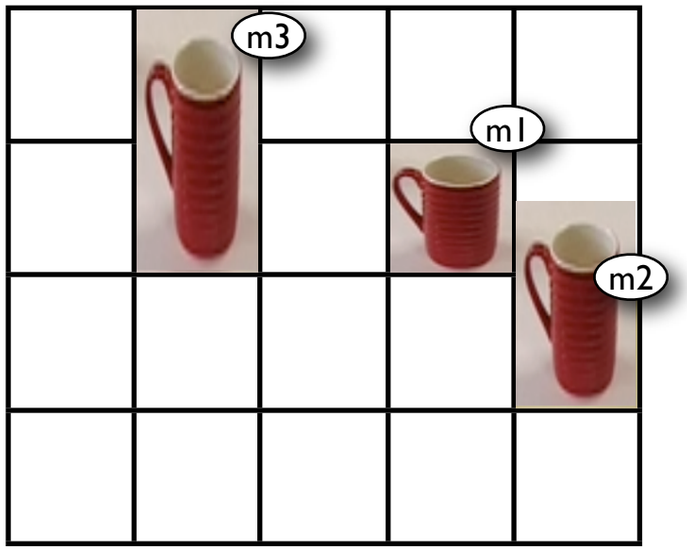
m2 Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	small	∅
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	m2

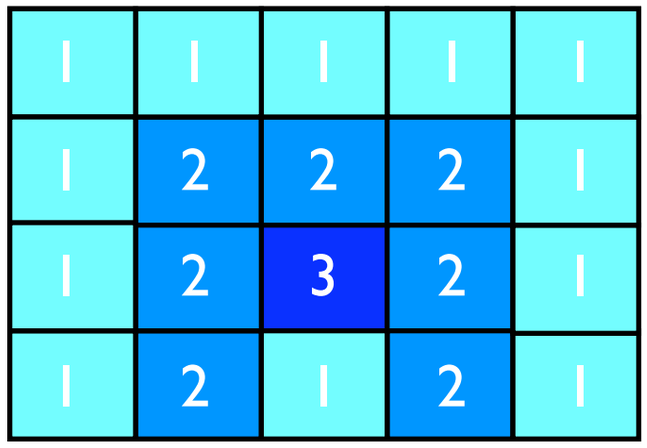


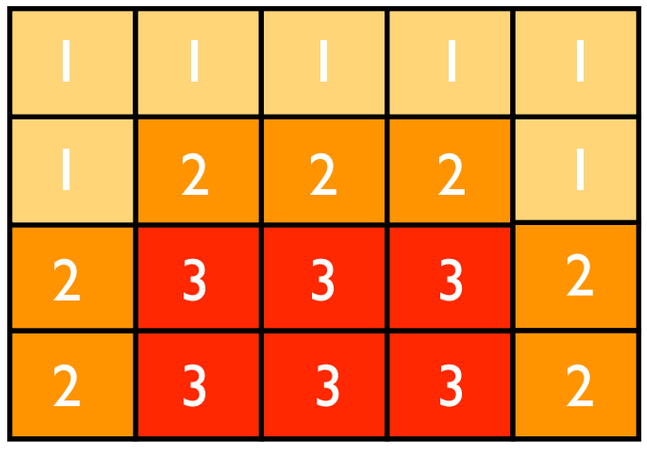


- Identification by contrastive material property
  - Distractors with same type but variation in contrastive properties
  - The individuals can be identified directly through contrast, either absolute or comparative (small, big; biggest, smallest; bigger, smaller)
  - E.g. “the big mug”, “the small mug”
  - (Additional distinctive properties: projective spatial relations)

Scene#CP.2 - data







m1

Characteristic	Value	Distractor
Type	mug	m2,m3
Color	red	m2,m3
Shape	contain	m2,m3
Size	smallest	∅
Topological		
ego	far	m3
allo-object	--	--
allo-human	near	m3
Projective		
ego	right	m2
allo-object	--	--
allo-human	left	m2
Vis.salience	2	m3
Kin. salience	2	m2,m3

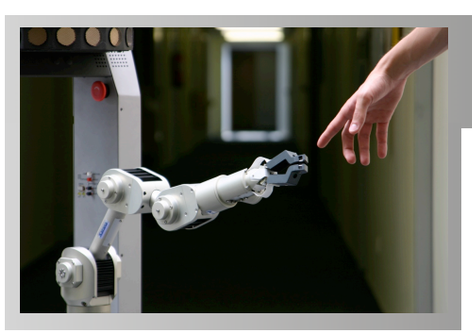
m2

Characteristic	Value	Distractors
Type	mug	m1,m3
Color	red	m1,m3
Shape	contain	m1,m3
Size	bigger	m2
Topological		
ego	near	∅
allo-object	--	--
allo-human	far	∅
Projective		
ego	right	m1
allo-object	--	--
allo-human	left	m1
Vis.salience	1	∅
Kin. salience	2	m1,m3

m3

Characteristic	Value	Distractors
Type	mug	m1,m2
Color	red	m1,m2
Shape	contain	m1,m2
Size	biggest	∅
Topological		
ego	far	m1
allo-object	--	--
allo-human	near	m1
Projective		
ego	left	∅
allo-object	--	--
allo-human	right	∅
Vis. salience	2	m1
Kin. salience	2	m1,m2

- Identification by contrastive property
  - Distractors with same type but variation in contrastive properties
  - The individuals can be identified directly through a combination of type, and contrastive property; sometimes an absolute use of a constrastive property is needed to identify - a relative, comparative use (smaller, bigger) will not work on this scene because there remain distractors then
  - (Additionally, non-object centered spatial relations are usually ambiguous, as are saliency measures)
  - (An interesting question is to what extent salience varies due to size here. E.g. compare m1 and m2. Currently, this is not taken into account)



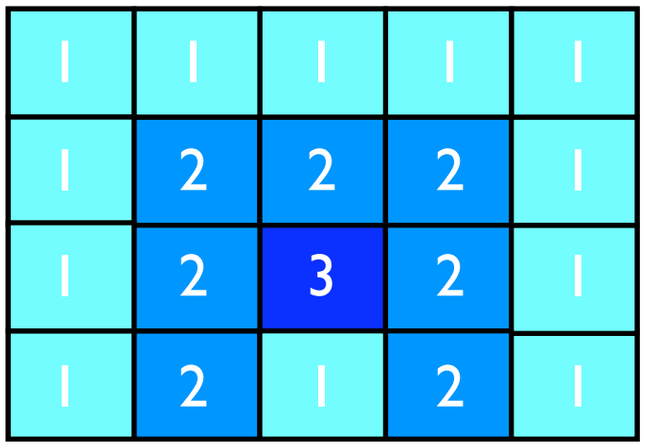
Identification by visual salience and type

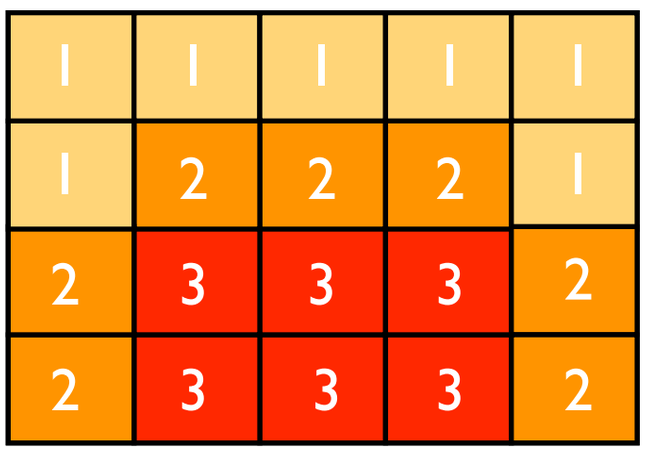
Scenes VST#1-

Scene#VST.1

m1		
Characteristic	Value	Distractors
Type	mug	m2
Color	white	∅
Shape	container	m2
Size	--	--
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	∅
Kinetic salience	3	∅

m2		
Characteristic	Value	Distractors
Type	mug	m1
Color	red	∅
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	1	∅
Kinetic salience	1	∅



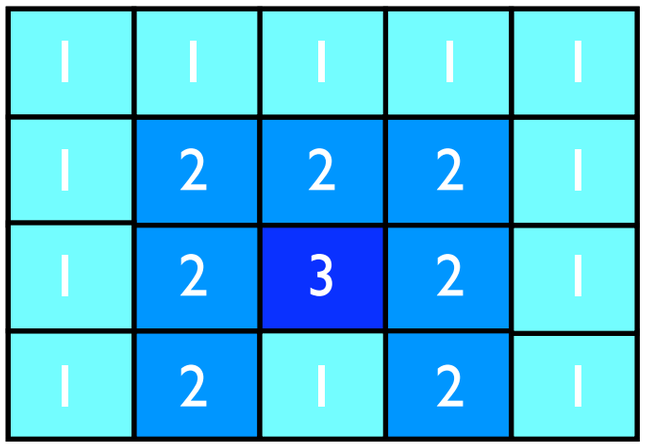


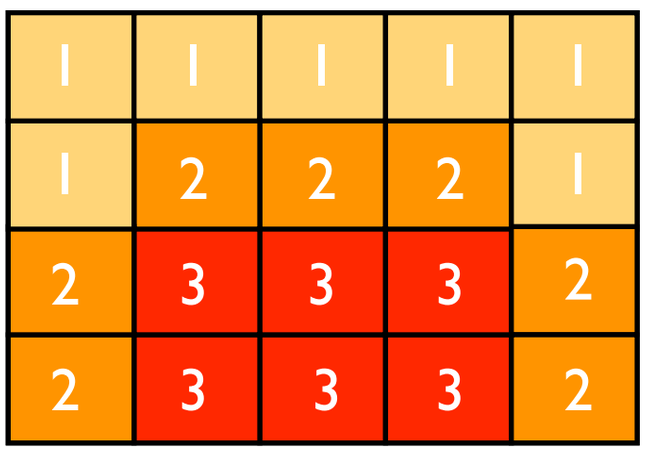
- Identification by distinctive visual salience
  - Distractors with same type but variation in visual salience
  - The individuals can be identified directly through type
  - E.g. “the mug” presumably refers to m2 (particularly in an action context)

Scene#VST.2

m1 Characteristic	Value	Distractors
Type	mug	m2
Color	white	∅
Shape	container	m2
Size	--	--
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	3	∅
Kinetic salience	3	∅

m2 Characteristic	Value	Distractors
Type	mug	m1
Color	red	∅
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	1	∅
Kinetic salience	1	∅



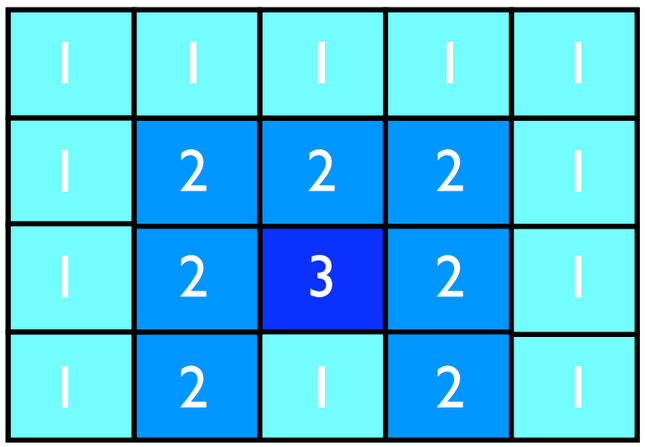


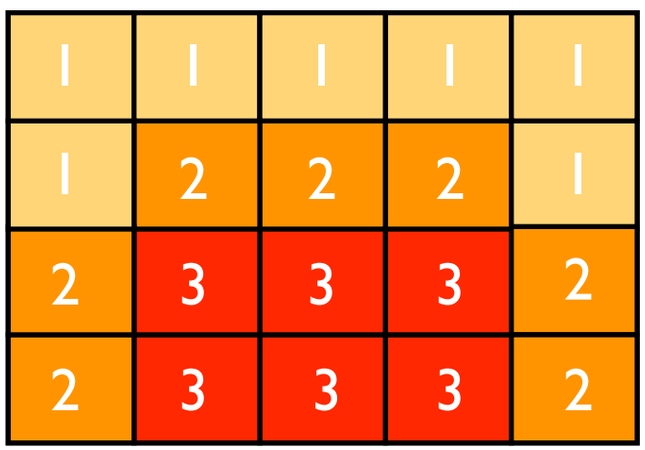
- Identification by distinctive visual salience
  - Distractors with same type but variation in visual salience
  - The individuals can be identified directly through type
  - The effect of "the mug" referring to m2 (particularly in an action context) is strengthened here (over VST#1) due to the increase in visual salience

Scene#VST.3

m1 Characteristic	Value	Distractors
Type	mug	m2
Color	white	∅
Shape	container	m2
Size	small	∅
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	3	∅
Kinetic salience	3	∅

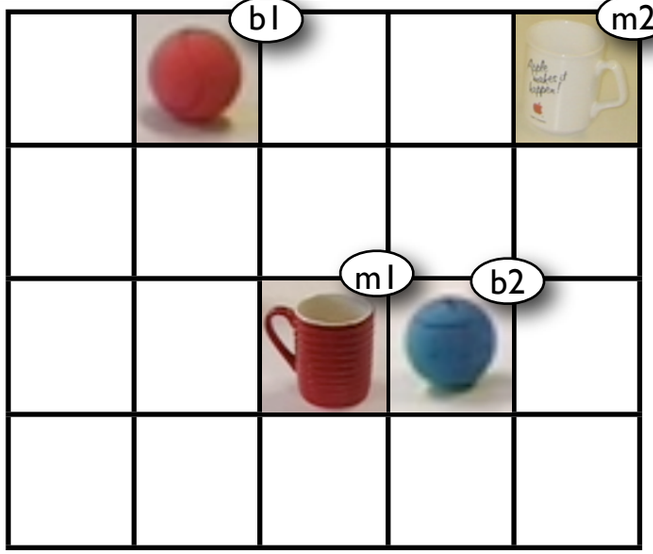
m2 Characteristic	Value	Distractors
Type	mug	m1
Color	red	∅
Shape	container	m1
Size	big(gest)	∅
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	1	∅
Kinetic salience	1	∅





- Identification by distinctive visual salience
  - Distractors with same type but variation in visual salience
  - The individuals can be identified directly through type
  - The effect of "the mug" referring to m2 (particularly in an action context) is potentially strengthened here (over VST#1,2) due to the increase in size

Scene#VST.4 - data

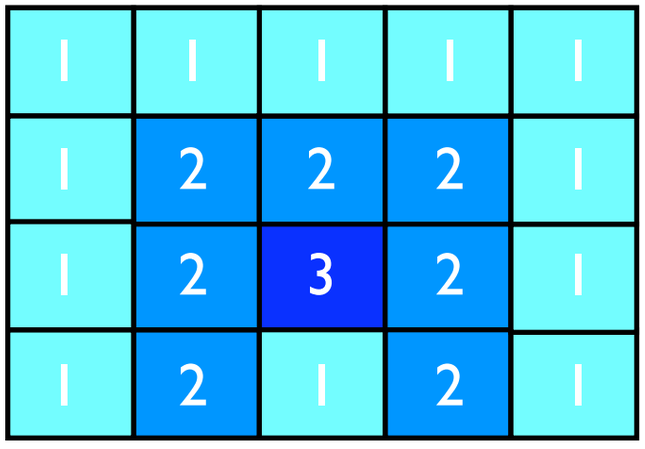


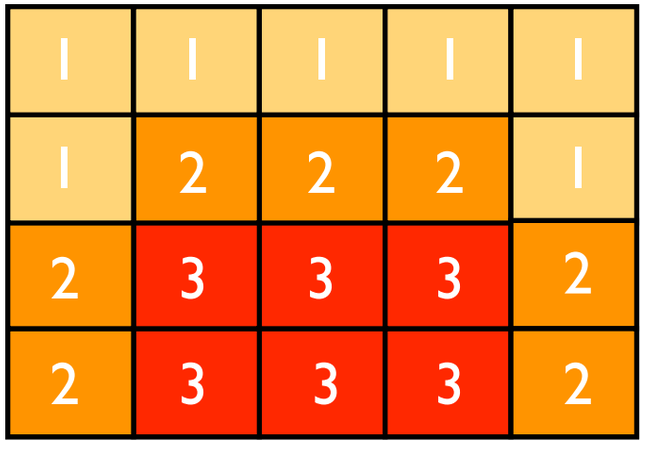
b1			
Characteristic	Value	Distractors	
Type	ball	b2	
Color	red	m1	
Shape	sphere	b2	
Size	--	--	
Topological			
ego	far	m2	
allo-object	--	--	
allo-human	near	m2	
Projective			
ego	left	m1	
allo-object	--	--	
allo-human	right	m1	
Vis. salience	1	m2	
Kin. salience	1	m2	

b2			
Characteristic	Value	Distractors	
Type	ball	b1	
Color	blue	∅	
Shape	sphere	b1	
Size	--	--	
Topological			
ego	near	m1	
allo-object	--	--	
allo-human	far	m1	
Projective			
ego	right	m2	
allo-object	--	--	
allo-human	left	m2	
Vis.salience	2	∅ (m1)	
Kin. salience	3	m1	

m1			
Characteristic	Value	Distractors	
Type	mug	m2	
Color	red	b1	
Shape	contain	m2	
Size	--	--	
Topological			
ego	near	b2	
allo-object	--	--	
allo-human	far	b2	
Projective			
ego	left	b1	
allo-object	--	--	
allo-human	right	b1	
Vis.salience	3	∅ (b2)	
Kin. salience	3	b2	

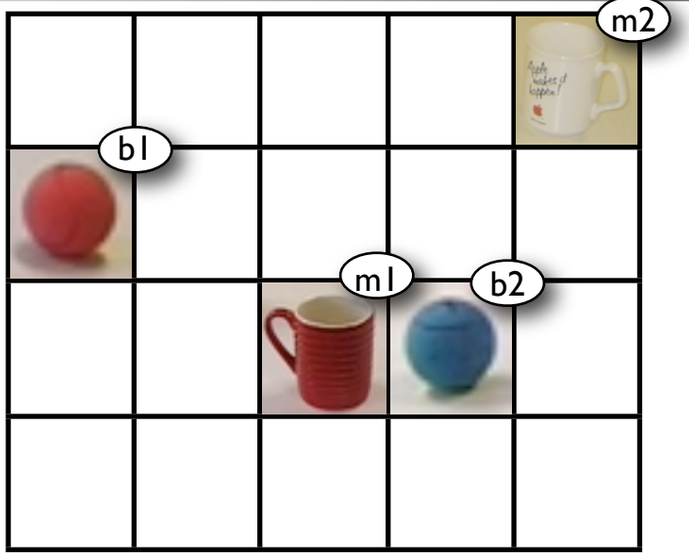
m2			
Characteristic	Value	Distractors	
Type	mug	m1	
Color	white	∅	
Shape	contain	m1	
Size	--	--	
Topological			
ego	far	b1	
allo-object	--	--	
allo-human	near	b1	
Projective			
ego	right	b2	
allo-object	--	--	
allo-human	left	b2	
Vis.salience	1	b1	
Kin. salience	1	b1	

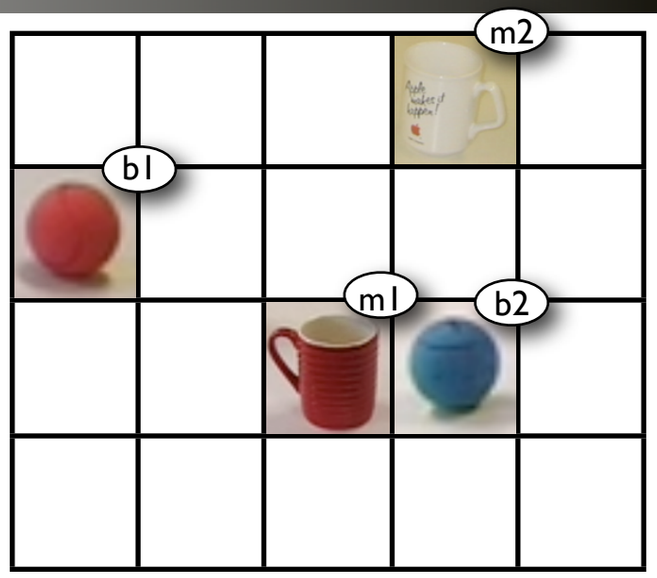


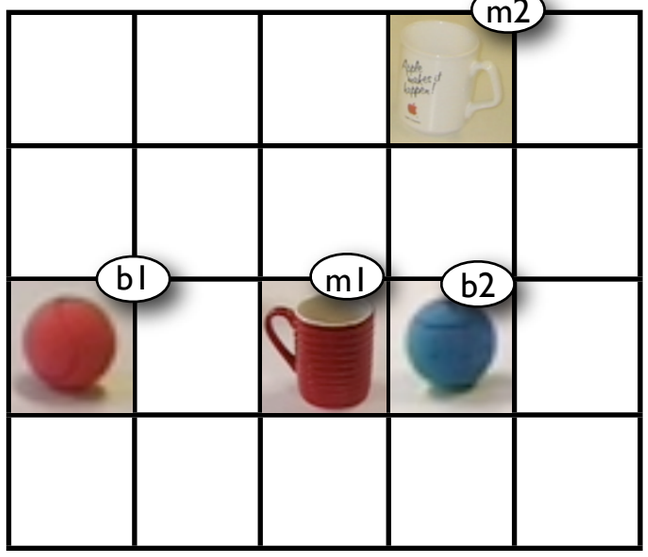


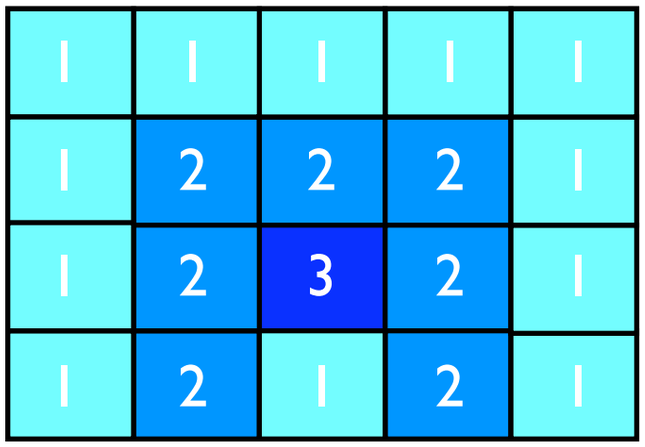
- Identification by distinctive visual salience
  - Distractors with same type but variation in visual salience
  - The individuals can be identified directly through type
  - E.g. “the mug” presumably refers to m1 (particularly in an action context), same with “the ball” referring to b2

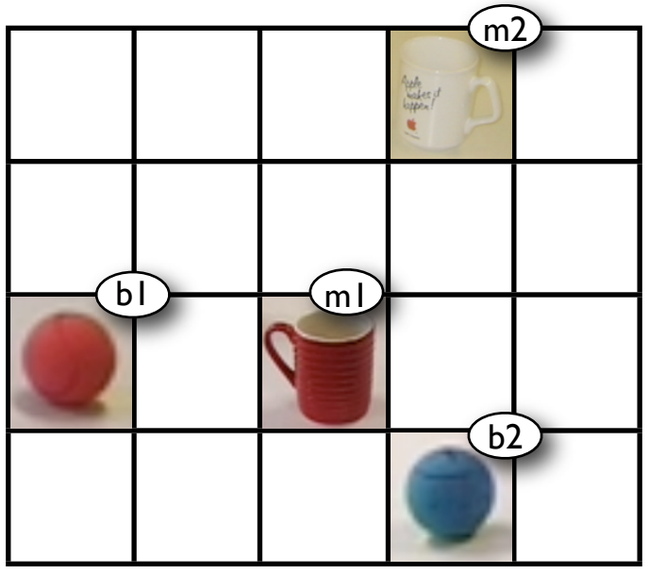
Scene#VST.4 - variations

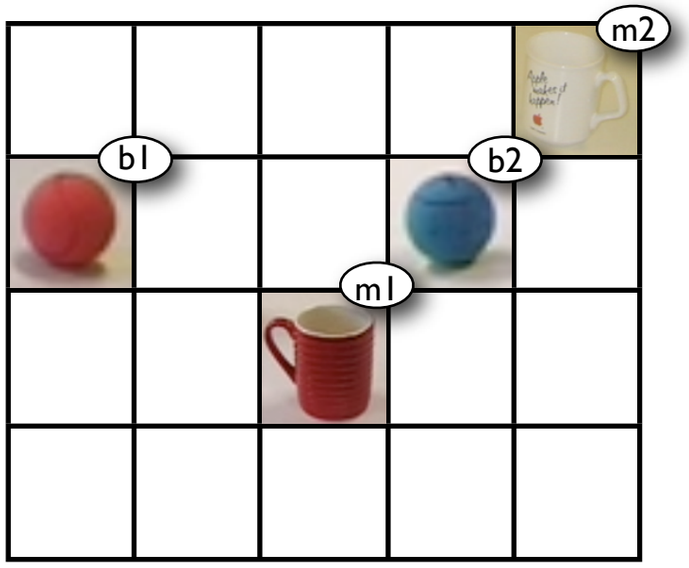


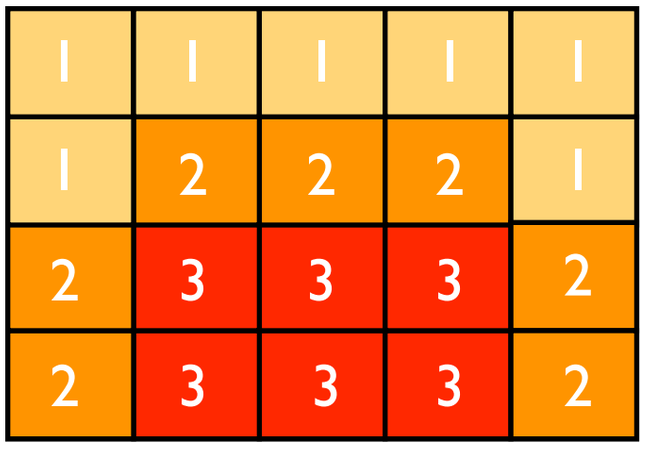




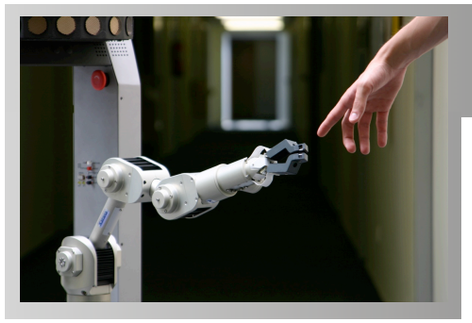








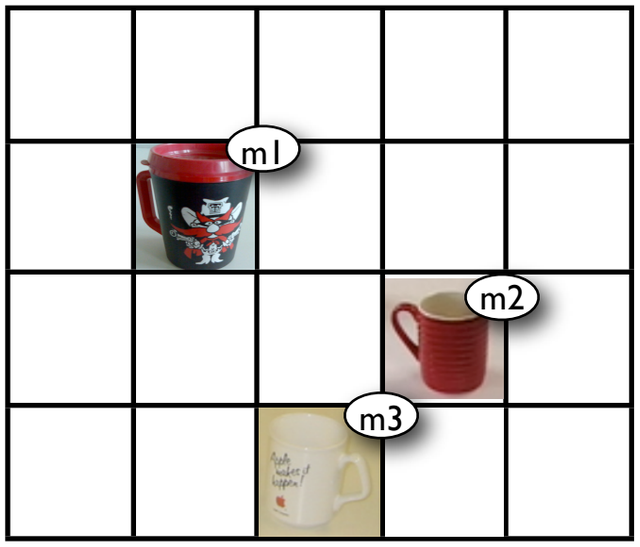
- Variations on VST.4 under equal visual salience patterns
  - Distractors as per VST.4, identical values for all characteristics
  - What changes is spatial placement (without changing relative spatial organization)

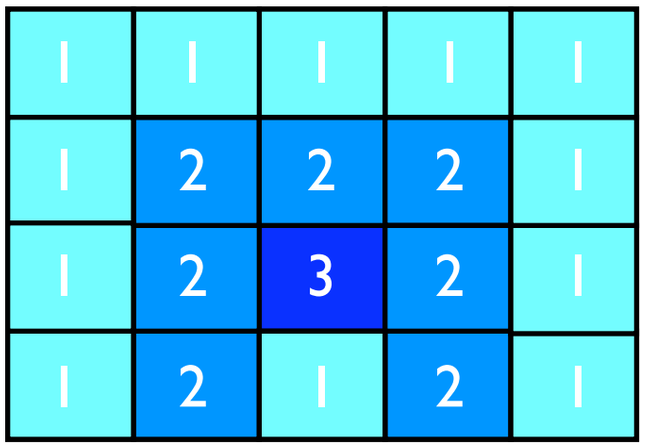


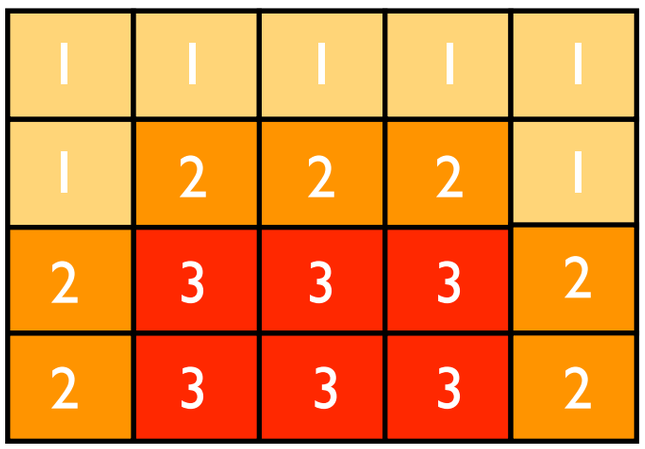

Identification by visual salience, type and material property

Scenes VSTCP#1-

Scene#VSTMP.1







m1

Characteristic	Value	Distractors
Type	mug	m2,m3
Color	red	m2,m3
Shape	container	m2,m3
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	2	∅

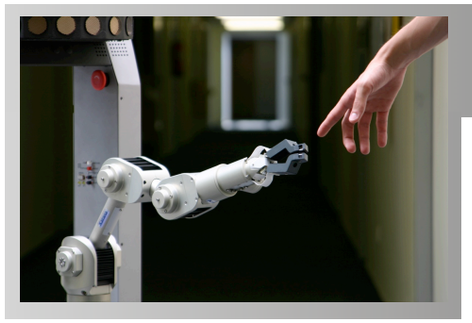
m2

Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	--	--
object-centered	--	--
human-centered	--	--
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	m3

m3

Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	∅ (m2)
object-centered	--	--
human-centered	far	∅ (m2)
Projective		
ego-centered	front	∅
object-centered	--	--
human-centered	--	--
Visual salience	1	∅
Kinetic salience	3	m2

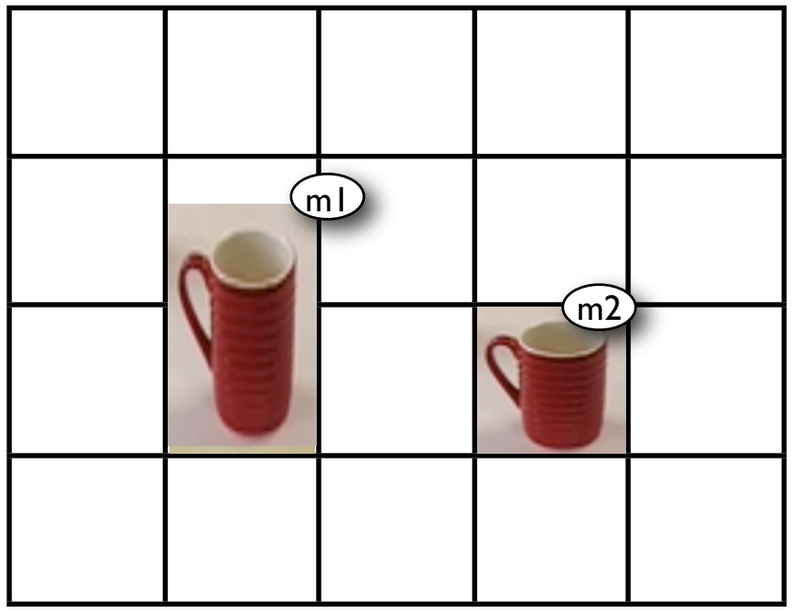
- Identification by distinctive combined salience
  - Distractors with same type but variation in combined salience
  - The individuals cannot be identified on visual or kinetic salience alone
  - To explain a (hypothesized) preference m2 for in the context of manipulation-related context requires a combination of visual and kinetic salience: m3 is just as reachable as m2, but as m3 is less visually salient than m2, m2 would be preferred for “take the mug”

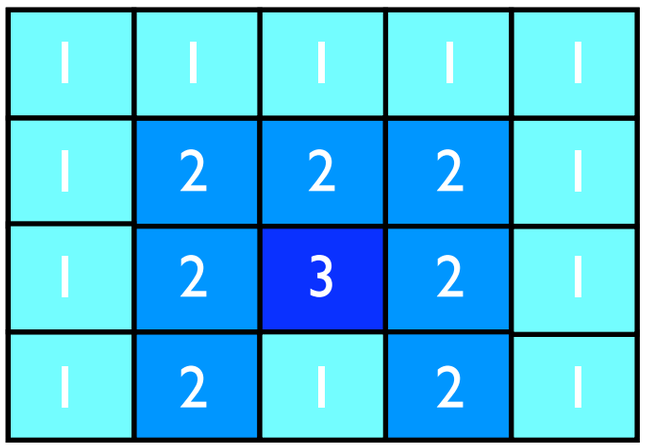


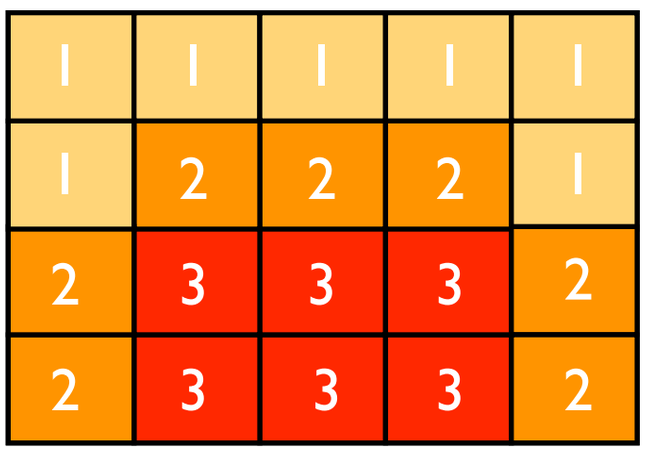

Identification by visual salience, type and contrastive property

Scenes VSTCP#1-

Scene#VSTCP.1







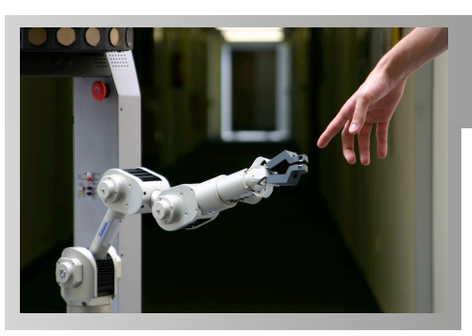
m1

Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	big	∅
Topological		
ego-centered	near	m2
object-centered	--	--
human-centered	far	m2
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	3	m2

m2

Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	small	∅
Topological		
ego-centered	near	m1
object-centered	--	--
human-centered	far	m1
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	m2

- Identification by distinctive material property
  - Distractors with same type but variation in contrastive properties
  - The individuals can be identified directly through contrast, either absolute or comparative (small, big; biggest, smallest; bigger, smaller)
  - E.g. “the big mug”, “the small mug”
  - (Additional distinctive properties: projective spatial relations)



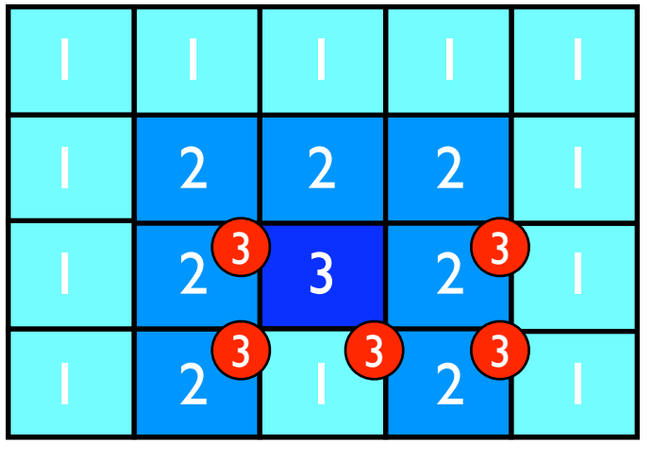
Identification by kinetic salience and type

Scenes KST#1-

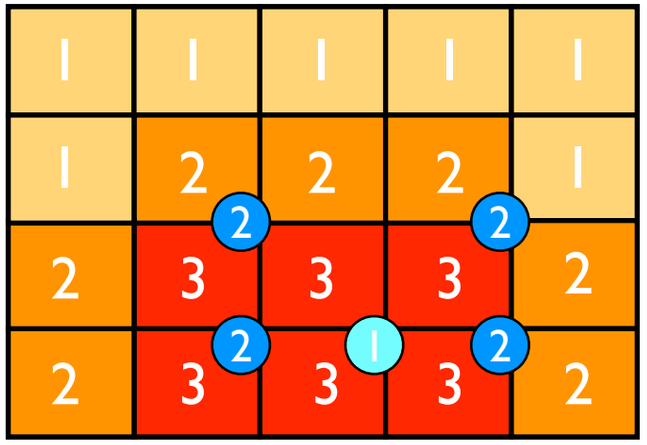
Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Between kinetic and visual salience

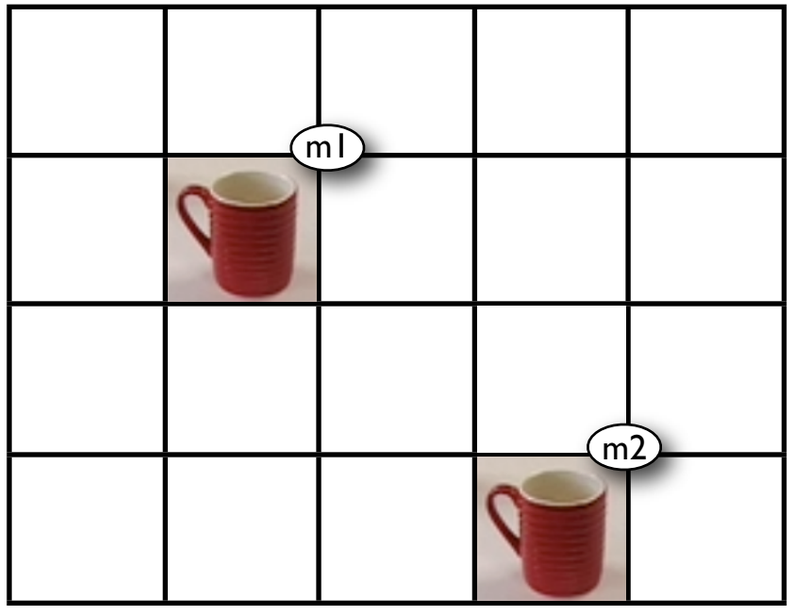
- Kinetic, or topokinetic, salience
  - Captures saliency of objects on the basis of reachability
  - This follows Berthoz' notion of *topokinetic* spatial understanding, maintained in parallel to a topological notion of space



- Between topokinetic and visual salience
  - We define topokinetic salience on area of reach, rather than visual focal center
  - This results in several cells having different values for kinetic and visual salience
  - As a consequence, distractor sets for kinetic and visual salience may be disjoint, or have non-empty disjoint complements of their intersection
  - We first specify scenes in which objects differ only in kinetic salience, i.e. under equal visual salience

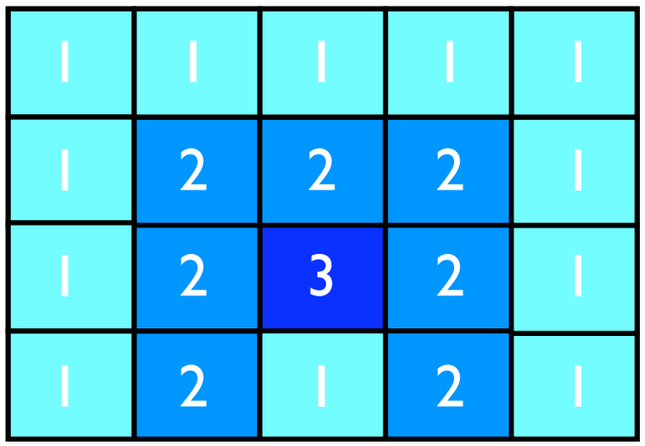


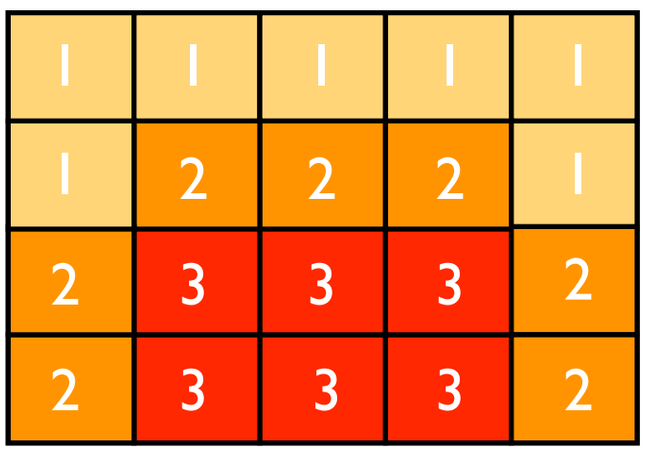
Scene#KST.1



m1		
Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	2	∅

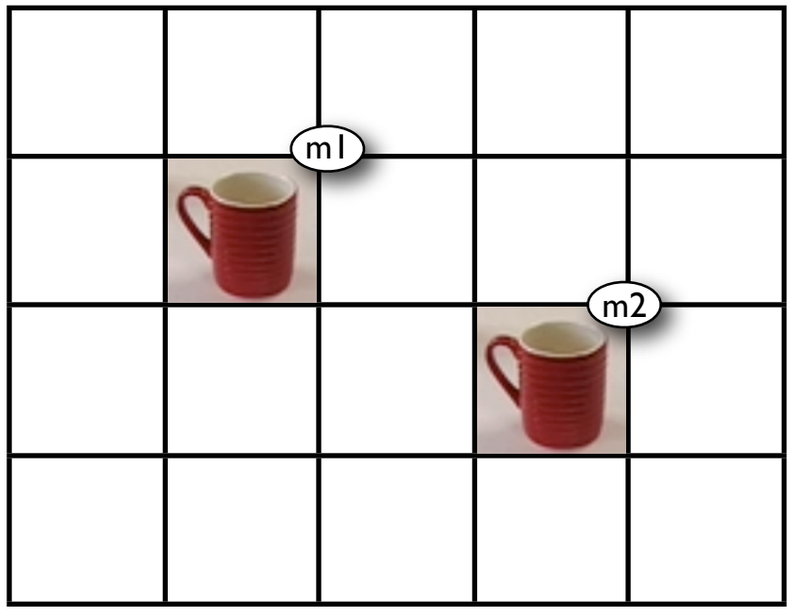
m2		
Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	∅





- Identification by distinctive kinetic salience
  - Distractors with same type but variation in kinetic salience
  - The individuals can be identified directly in the context of manipulation-related context
  - E.g. “take the mug” will prefer taking m2, not m1

Scene#KST.2



1	1	1	1	1
1	2	2	2	1
1	2	3	2	1
1	2	1	2	1

1	1	1	1	1
1	2	2	2	1
2	3	3	3	2
2	3	3	3	2

m1

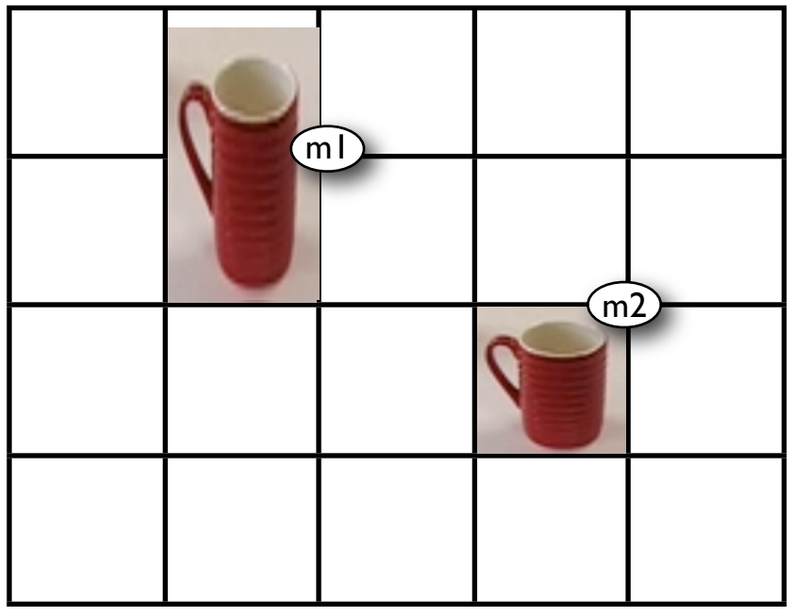
Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	2	∅

m2

Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	∅

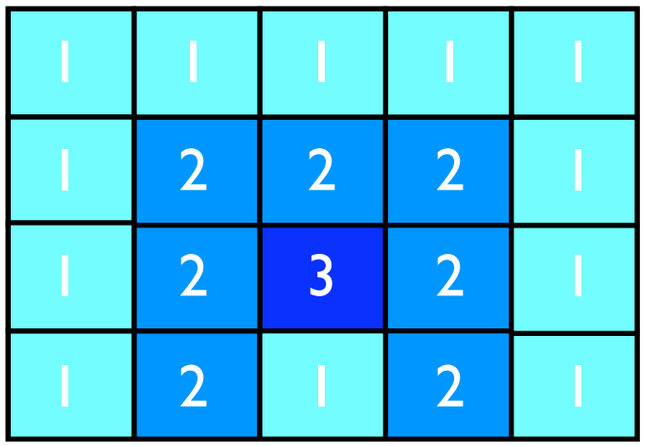
- Identification by distinctive kinetic salience
  - Distractors with same type but variation in kinetic salience
  - The individuals can be identified directly in the context of manipulation-related context
  - E.g. “take the mug” should prefer taking m2, not m1
  - Interesting is, though, whether the difference in visual salience and kinetic salience is enough (when considering them as continuous functions, rather than discrete functions)

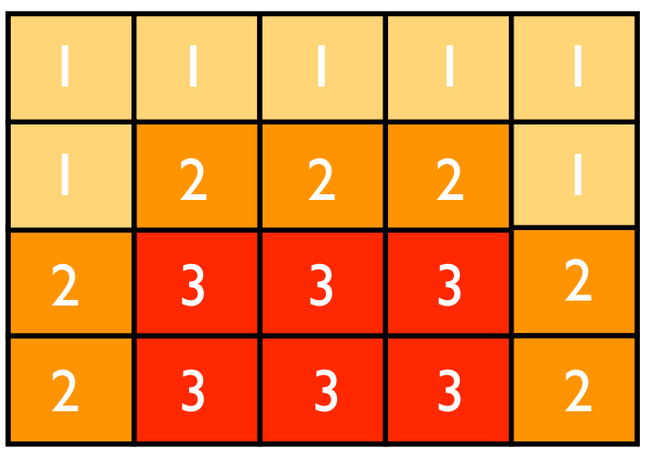
Scene#KST.2 - variants



m1		
Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	big(ger)	∅
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	2	∅

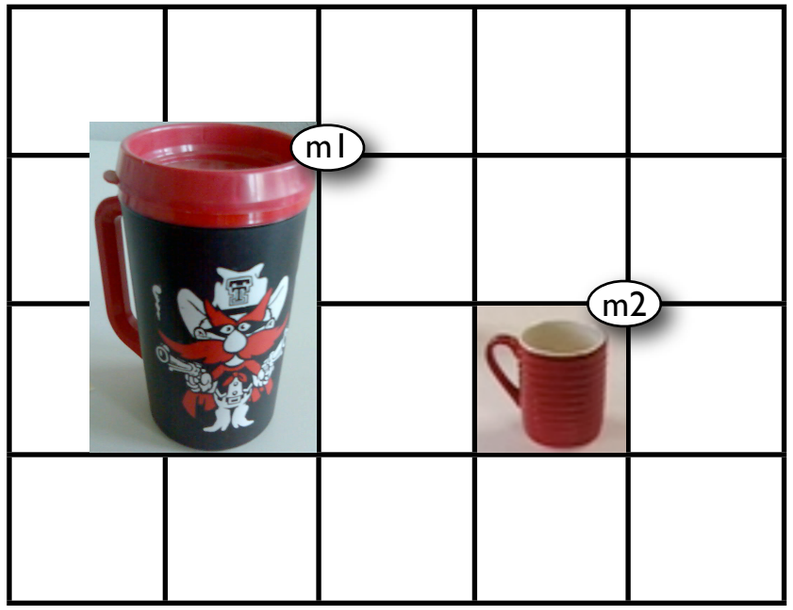
m2		
Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	small(er)	∅
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	∅





- Identification by distinctive kinetic salience
  - Distractors with same type but variation in kinetic salience
  - Interesting is whether the difference in visual salience and kinetic salience is enough (considering them as continuous functions), particularly if the kinetically less salient *may be* visually more salient due to size

Scene#KST.3



1	1	1	1	1
1	2	2	2	1
1	2	3	2	1
1	2	1	2	1

1	1	1	1	1
1	2	2	2	1
2	1	3	3	2
2	3	3	3	2

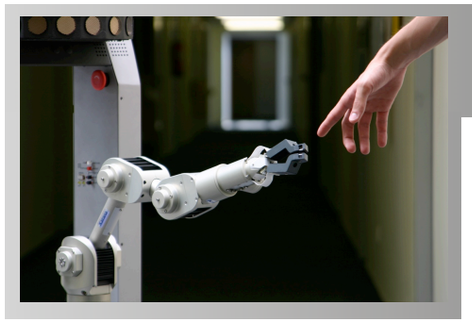
m1

Characteristic	Value	Distractors
Type	mug	m2
Color	red	m2
Shape	container	m2
Size	big(ger)	∅
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	1	∅

m2

Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	small(er)	∅
Topological		
ego-centered	near	∅
object-centered	--	--
human-centered	far	∅
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	∅

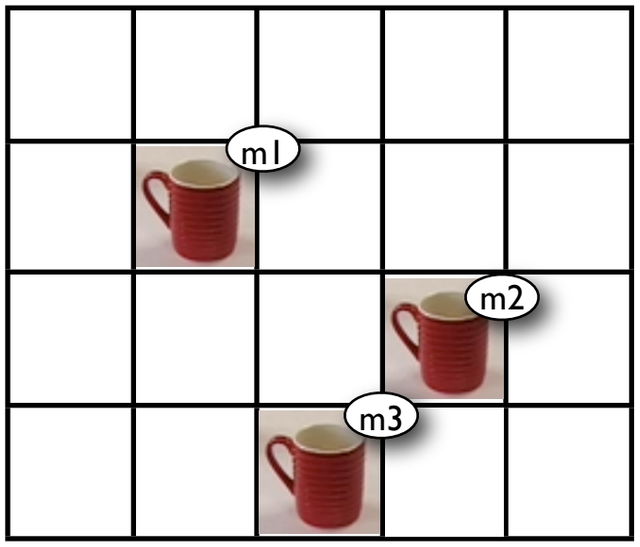
- Identification by distinctive kinetic salience
  - Distractors with same type but variation in kinetic salience due to differing graspability of the object
  - Assuming that m1 is not graspable for the system, its salience would (presumably) be low - although reachable, it isn't graspable, and so for "Take the mug" m2 would be the preferred referent
  - This raises the question to what extent graspability should be taken into account in kinetic salience - whether this is an abstract criterion, or action-dependent: for "push the mug" the situation would presumably be ambiguous, as kinetic salience wouldn't need to be changed as graspability is not required.



Identification by combined salience and type

Scenes CST#1-

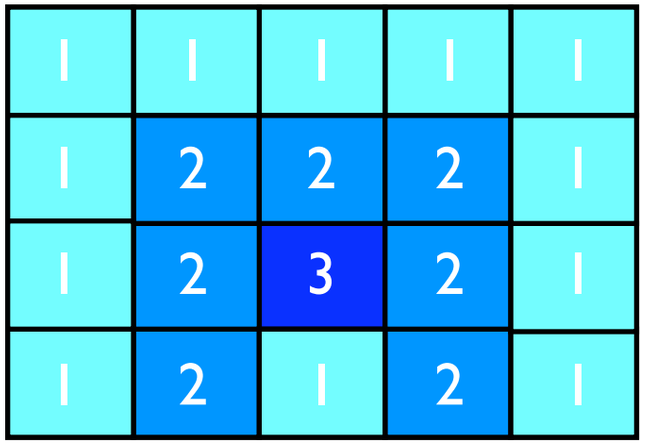
Scene#CST.1

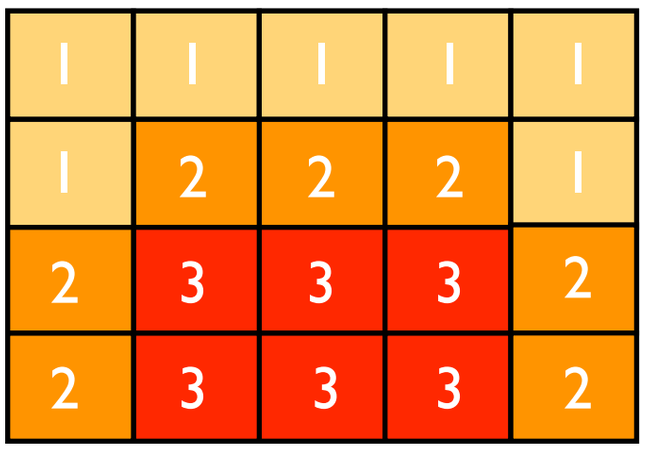


m1 Characteristic	Value	Distractors
Type	mug	m2,m3
Color	red	m2,m3
Shape	container	m2,m3
Size	--	--
Topological		
ego-centered	far	∅
object-centered	--	--
human-centered	near	∅
Projective		
ego-centered	left	∅
object-centered	--	--
human-centered	right	∅
Visual salience	2	m2
Kinetic salience	2	∅

m2 Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	--	--
object-centered	--	--
human-centered	--	--
Projective		
ego-centered	right	∅
object-centered	--	--
human-centered	left	∅
Visual salience	2	m1
Kinetic salience	3	m3

m3 Characteristic	Value	Distractors
Type	mug	m1
Color	red	m1
Shape	container	m1
Size	--	--
Topological		
ego-centered	near	∅ (m2)
object-centered	--	--
human-centered	far	∅ (m2)
Projective		
ego-centered	front	∅
object-centered	--	--
human-centered	--	--
Visual salience	1	∅
Kinetic salience	3	m2





- Identification by distinctive combined salience
  - Distractors with same type but variation in combined salience
  - The individuals cannot be identified on visual or kinetic salience alone
  - To explain a (hypothesized) preference m2 for in the context of manipulation-related context requires a combination of visual and kinetic salience: m3 is just as reachable as m2, but as m3 is less visually salient than m2, m2 would be preferred for “take the mug”