



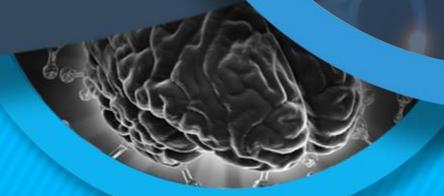
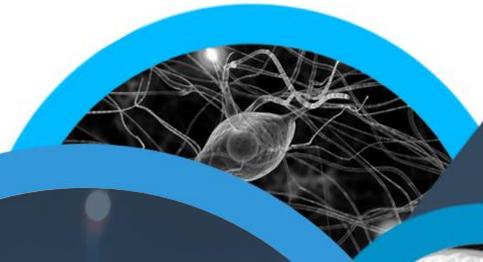
COGNITIVE SCIENCE
IN SEARCH OF UNITY



NATIONAL SCIENCE CENTRE

Marcin Miłkowski
NCN SONATA BIS Grant #14/E/HS1/00803

Cognitive Architectures and the Unity of Cognitive Science



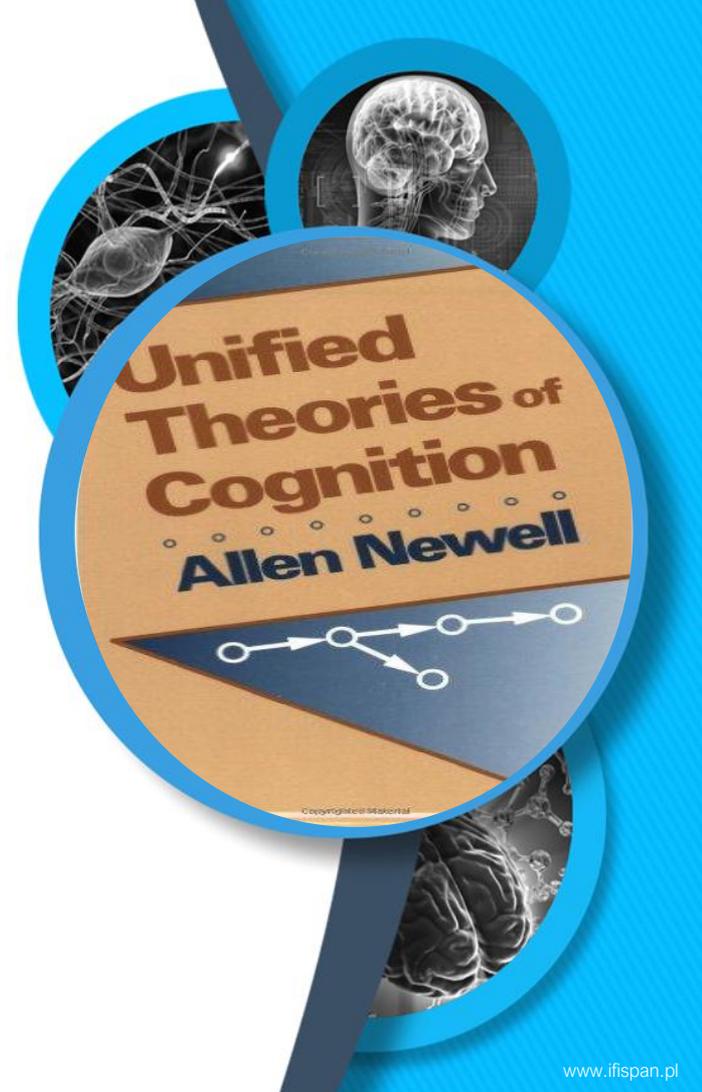
Plan of the talk

- Unification in cognitive science: a primer
- ... and a graded multidimensional account
- Cognitive architectures and unified theories
- Unified theories vs. integrated architectures
- Global and local understanding

Unification in Cognitive Science

Allen Newell (1973) argued: you can't play 20 questions with nature and win!

- Single tasks and micro-theories
- No unity in explanations
- Need for an overarching theory of cognition
- Newell's proposal: study the architecture of cognitive systems, and build unified theories by appeal to architectures



Unification: a graded account

Unified explanations are not monstrous.
Uniformity of explanations may be analyzed as:

- Simplicity, or non-redundancy (Akaike information criterion, Solomonoff-Kolmogorov complexity)
- Invariance, or generality (scope of the explanation)
- Elegance, or beauty (non-monstrosity)
- Systematicity



Unification: a graded account

Unification can be graded (as measures of scope / invariance, simplicity, homogeneity are): these are four dimensions of a complex achievement

- **This can be applied to theories, models, and to research traditions:**
 - Miłkowski, M., & Nowakowski, P. (2019). Representational unification in cognitive science: Is embodied cognition a unifying perspective? *Synthese*. <https://doi.org/10.1007/s11229-019-02445-w>

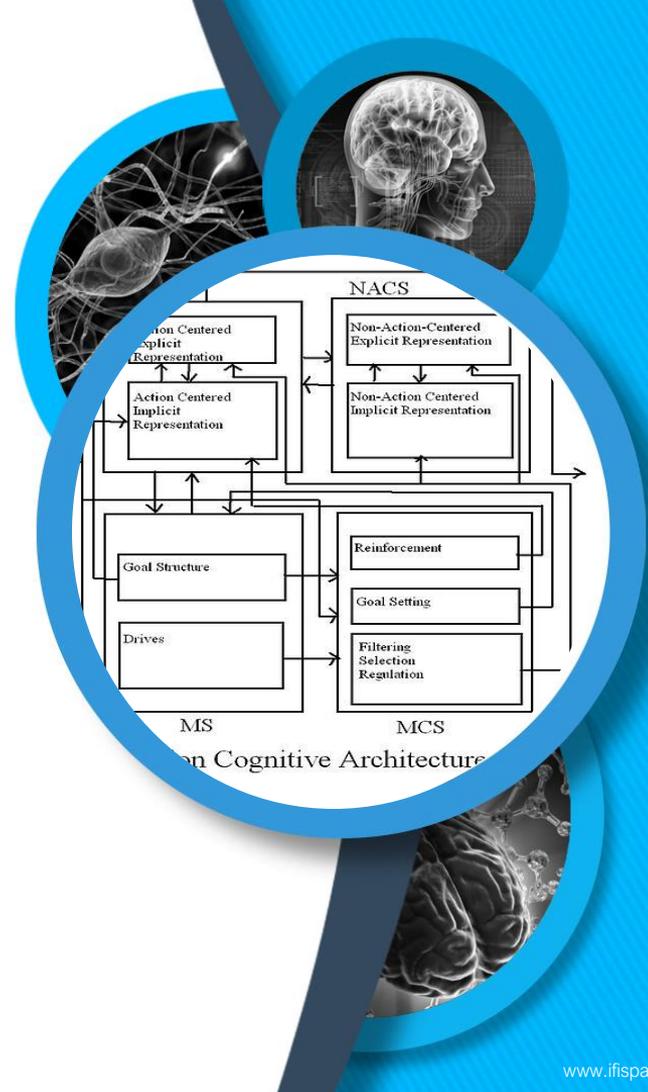


Cognitive architectures

- Build a general outline for a cognitive system
- Tempting: Rephrase existing models to match the architecture

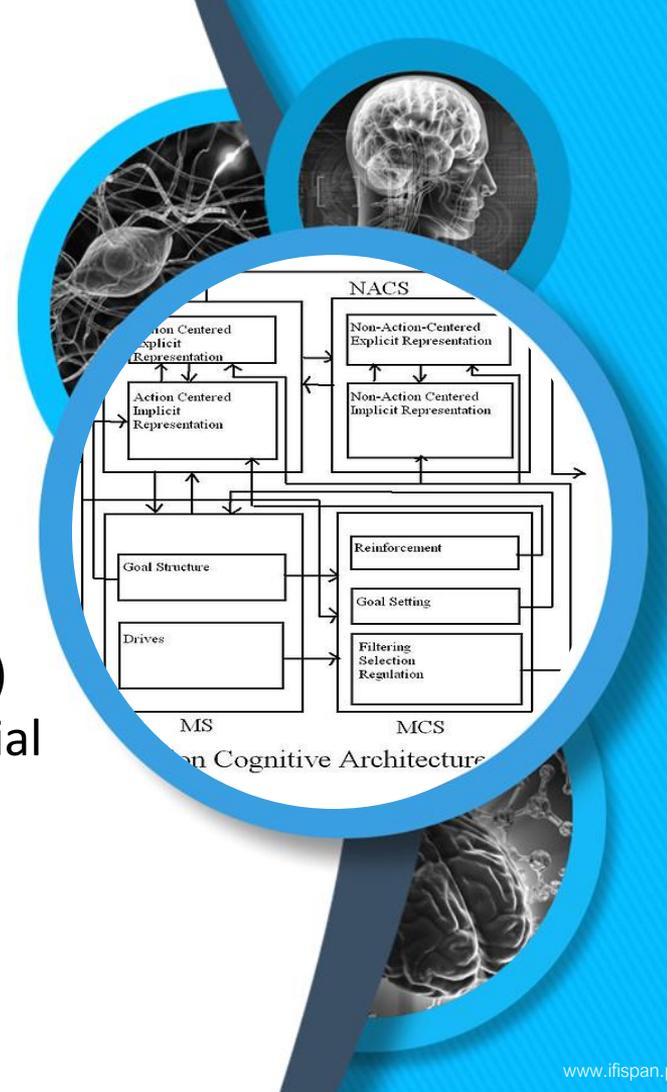
Tradeoffs:

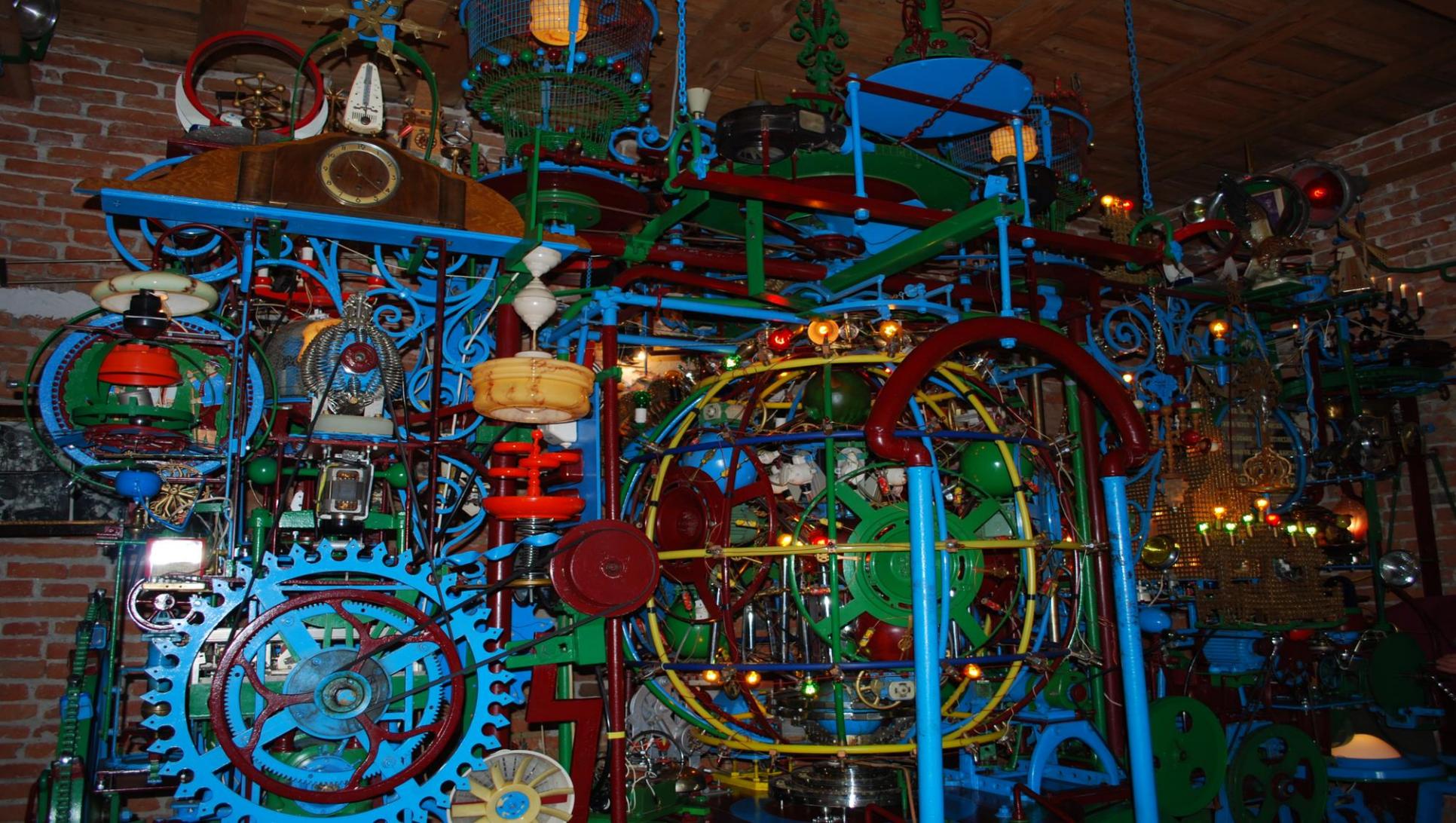
- Simplifications
- Proofs of concept or partial models instead of end-to-end modeling
- Superficial data fitting
- Possible monstrous models (aka modeljunk)



Can architectures provide unified theories?

- Simplicity, or non-redundancy
 - Single mechanism / blueprint
- Systematicity
 - The blueprint + levels of abstraction / explanation (also bands, scales etc.)
- Invariance, or generality (scope of the explanation)
 - Avoid shallow criticism but note the lack of social interactions, external artifacts in distributed cognition, or aspects of consciousness
- Elegance, or beauty (non-monstrosity)





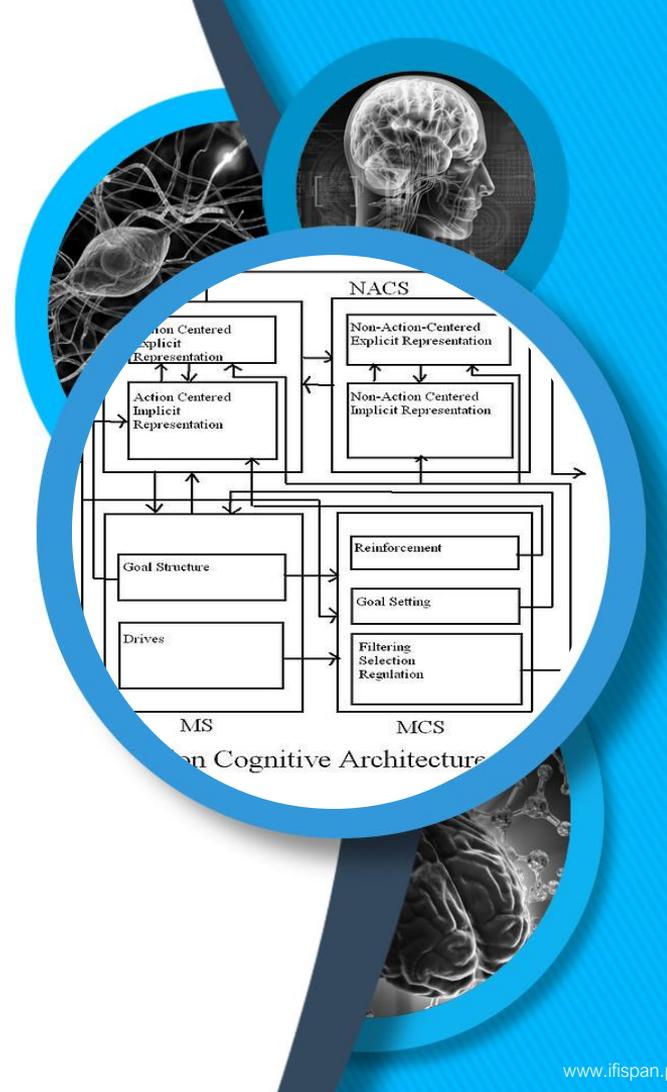
Monstrosity

- Not all coherent representations are good scientific theories.
- This is why disunified but integrated representations may be monsters.
- The more monstrous the representation, the less it is unified.

Votsis 2015:

Monstrosity = level of confirmational disconnectedness

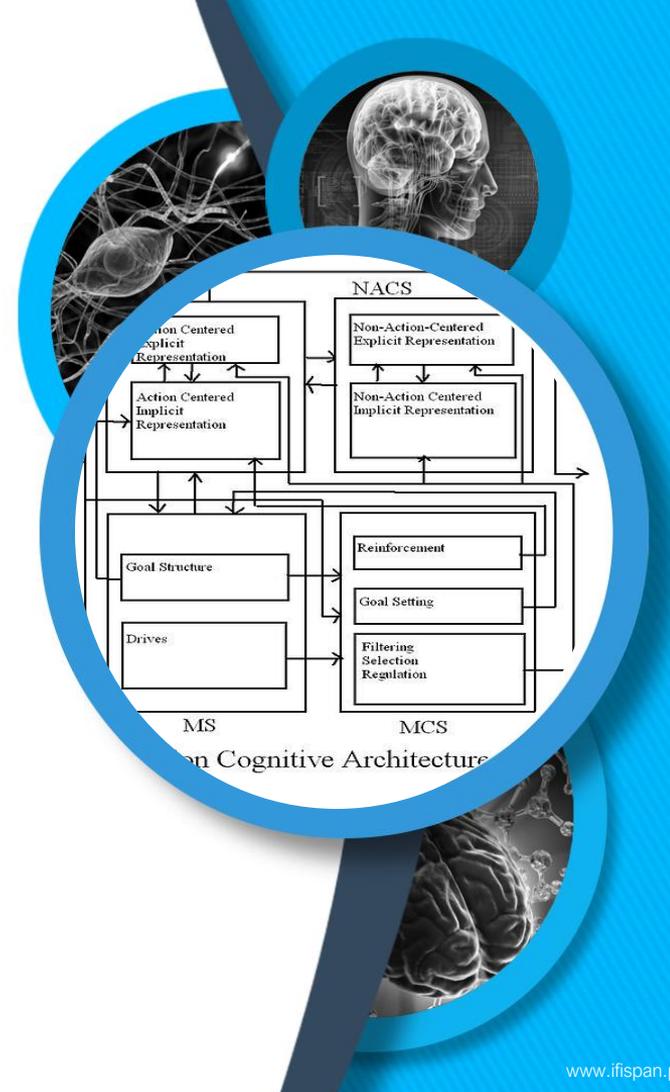
Any two content parts of a non-self-contradictory proposition Γ expressed as propositions A, B are *confirmationally disconnected* if, and only if, for all pairs of internally non-superfluous propositions α, β where $A \vdash_r \alpha$ and $B \vdash_r \beta$: (i) there is no true or partly true proposition γ such that $\alpha \vdash_r \gamma$ and $\beta \vdash_r \gamma$ and (ii) where $0 < P(\alpha), P(\beta) < 1, P(\alpha/\beta) = P(\alpha)$ and (iii) there is no atomic proposition δ such that $\alpha \wedge \beta \vdash_r \delta, \alpha \not\vdash_r \delta$ and $\beta \not\vdash_r \delta$.



Monstrosity

Example:

ACT-R model of past tense acquisition (Taatgen & Anderson 2002): the detail of the model (e.g., how the past form is retrieved by analogy) may be rejected without rejecting the rational analysis of the phenomenon or ACT-R architecture.



Integration as constraining

Scientific representations are integrated with other representations as long as they are constrained.

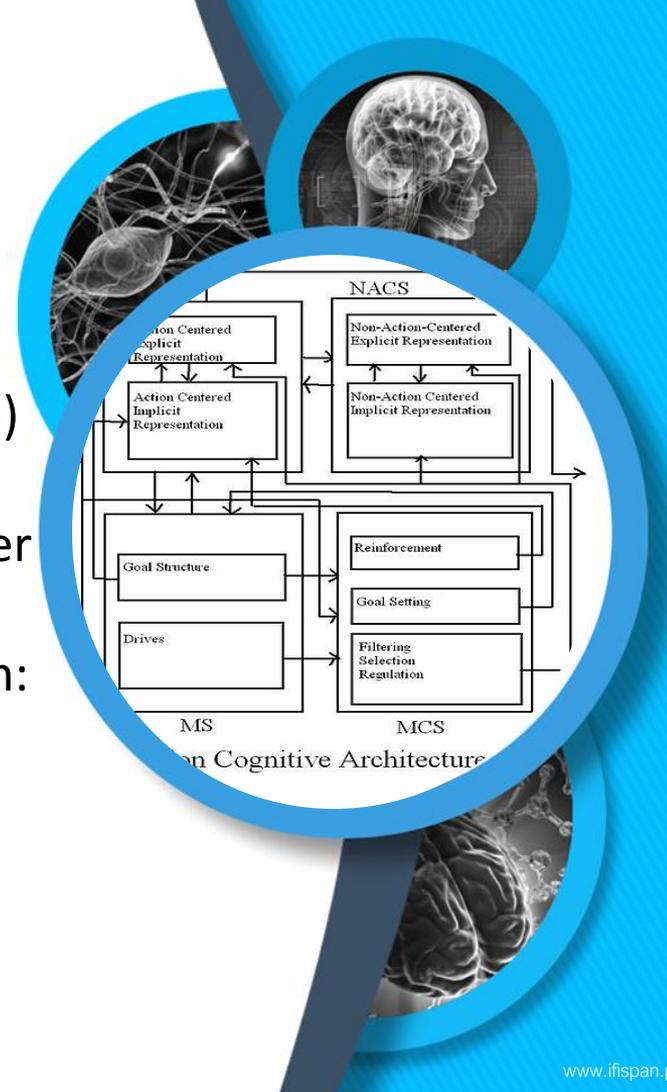
- Constraints may be symmetrical or not (consider fundamental physics and electric chemistry).
- Simplest constraint: truth-value, or logical coherence
- Mechanistic constraints:
 - Same entities & activities referred to
 - Spatial and temporal features
- In general: constraint satisfaction = coherence of integrated representations





Unified theories vs. architectures

- Architectures (computationally defined) are not to be confused with theories (Cooper & Shallice 1995)
- But theories require going beyond mere effect fishing to bigger principles (witness excitement over principles such as the FEP)
- Informative unifying principles should provide both:
 - Global understanding of how things hang together
 - Local understanding of particular phenomena (Bangu 2017)



Local and global understanding

- While grand principles are easy for global understanding (take the first principle of cognitive science), these provide little predictive insights about particular phenomena (cf. Litwin & Miłkowski 2020).
- Checklist-style ‘Newell tests’ (or Norman’s 12 issues [1981]) are a way to assess systematicity and generality (ability)
- But we might try to tease out dimensions over which cognitive mechanisms could be invariant to see whether principles are informative for them...

The main principle of cognitive science:

Cognition requires computation and cognitive representation for the adaptive control behavior to occur.

Conclusions

Unified theories of cognition or mind are difficult to produce.

- But they provide both global and local understanding, whose value is difficult to undermine.
- Systematic theories of how cognitive mechanisms could vary are likely to be a way forward at least for achieving more rounded 'Newell tests'.



NATIONAL SCIENCE CENTRE

NCN SONATA BIS Grant #14/E/HS1/00803

Cognitive Science In Search of Unity

Website : <http://cognitive.ifispan.pl>

Email : marcin.milkowski@gmail.com

- Miłkowski, M. (2016). Unification Strategies in Cognitive Science. *Studies in Logic, Grammar and Rhetoric*, 48(1), 13–33. <https://doi.org/10.1515/slgr-2016-0053>
- Miłkowski, M., & Nowakowski, P. (2019). Representational unification in cognitive science: Is embodied cognition a unifying perspective? *Synthese*. <https://doi.org/10.1007/s11229-019-02445-w>
- Miłkowski, M. (2016). Integrating cognitive (neuro)science using mechanisms. *Avant: Journal of Philosophical-Interdisciplinary Vanguard*, VI(2), 45–67. <https://doi.org/10.26913/70202016.0112.0003>
- Litwin, P., & Miłkowski, M. (2020). Unification by Fiat: Arrested Development of Predictive Processing. *Cognitive Science*, 44(7), e12867. <https://doi.org/10.1111/cogs.12867>