



**SucFoDok**  
Suchtpräventionsforschung  
und -dokumentation

**KLINIKUM  
AKADEMIE  
FORSCHUNG**



**ANTON-PROKSCH-INSTITUT**

# 'Evaluation' From Dogma to Useful Tool



**Alfred Uhl**

Uhl (2007): How to Camouflage Ethical Questions in Addiction Research.  
In: Fountain, J.; Korf, D.J. (eds.): Drugs in Society European Perspectives. Radcliff, Oxford

MONOGRAFIE

N° 13

Alfred Uhl



Jak maskovat etické otázky  
ve výzkumu závislostí

Jestliže je myšlení příjemné, pak ještě nezačalo

NÁRODNÍ MONITOROVACÍ STŘEDISKO PRO DROGY A DROGOVÉ ZÁVISLOSTI

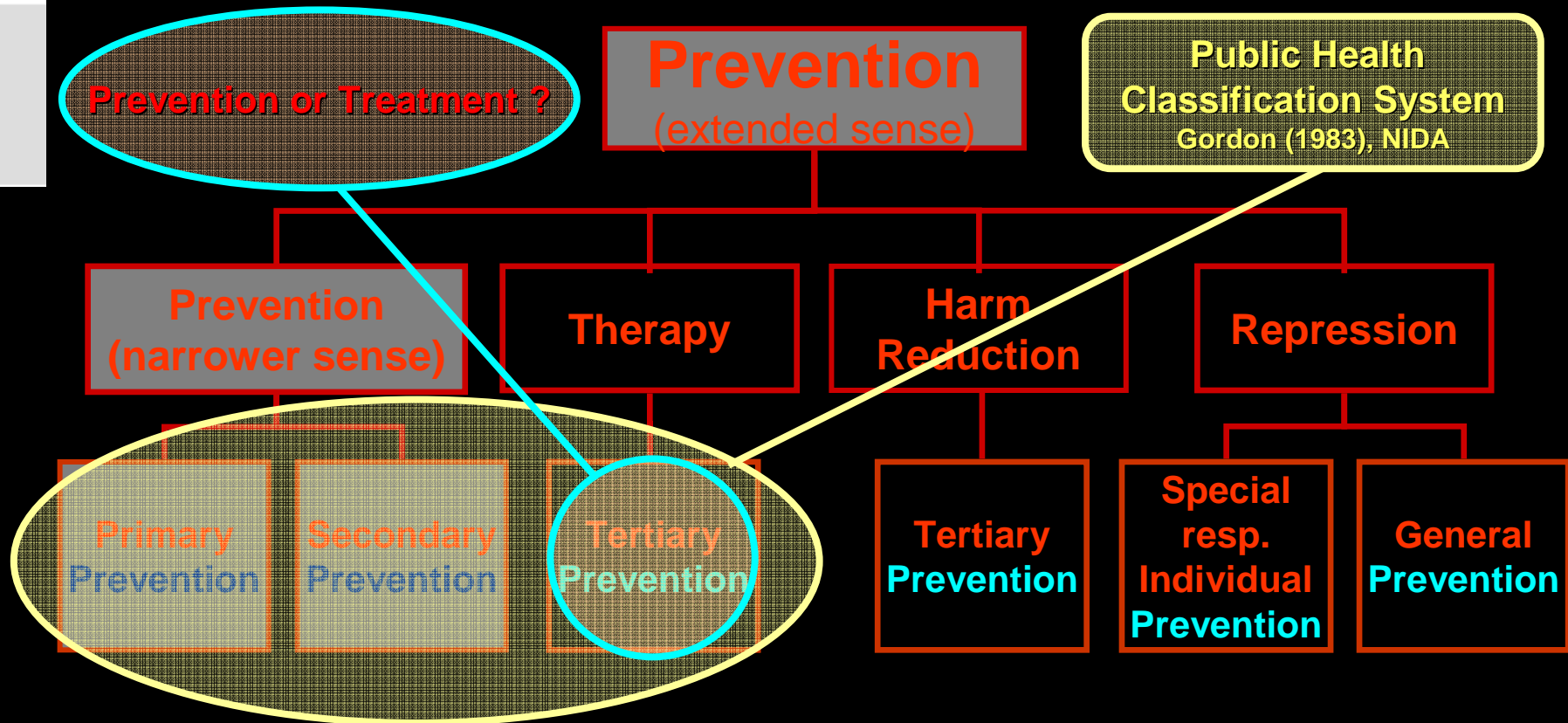


**(1)**

**What is  
„Substance Abuse Prevention“ ?**

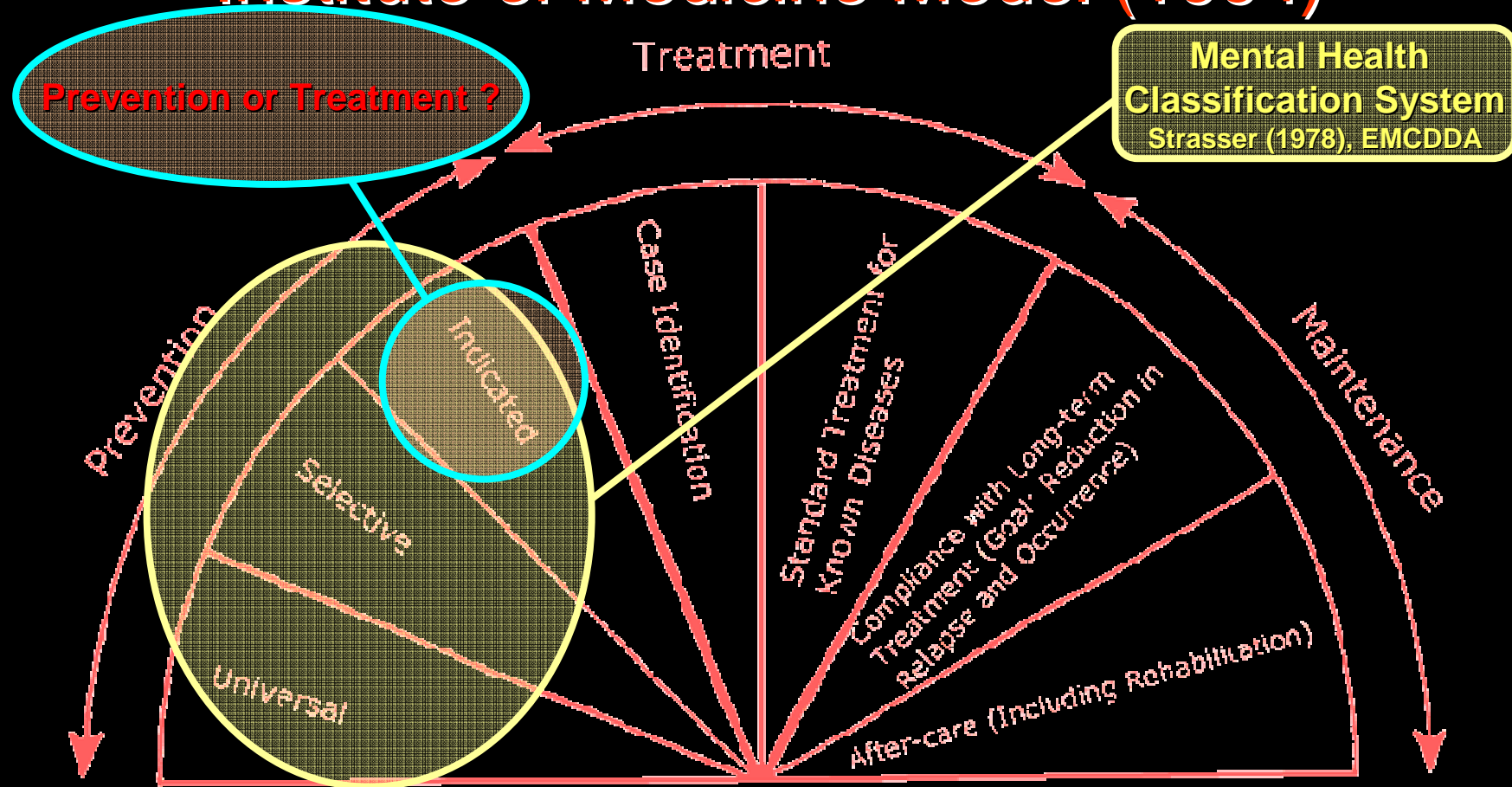
# Classification of Prevention

## Four-Column-Model



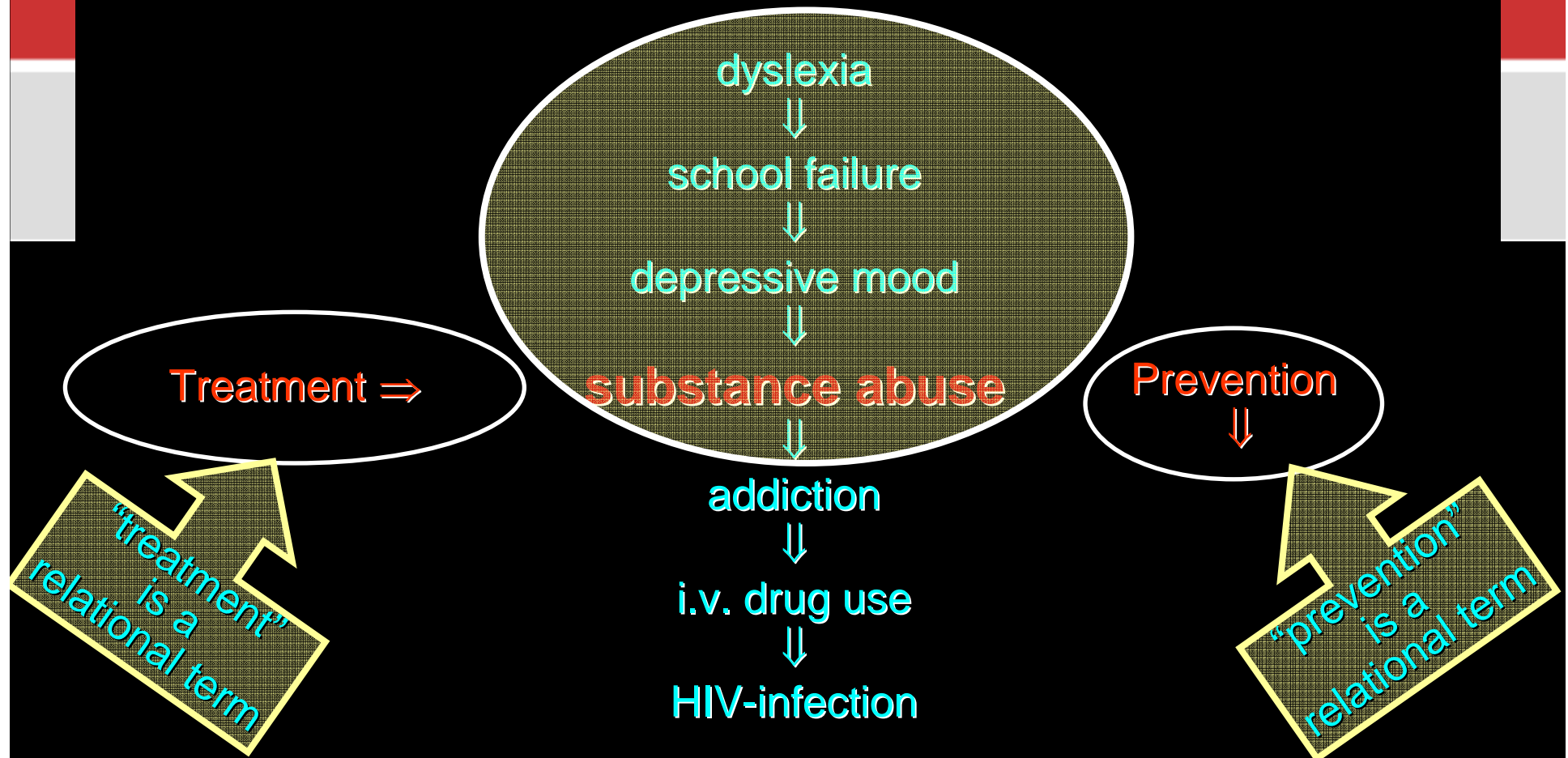
Uhl, A.; Gruber, Ch. (2004): Suchtprävention. In: Brosch, R.; Mader, R. (Hrsg.): Sucht and Suchtbehandlung: Problematik and Therapie in Österreich. LexisNexis, Wien

# Institute of Medicine Model (1994)

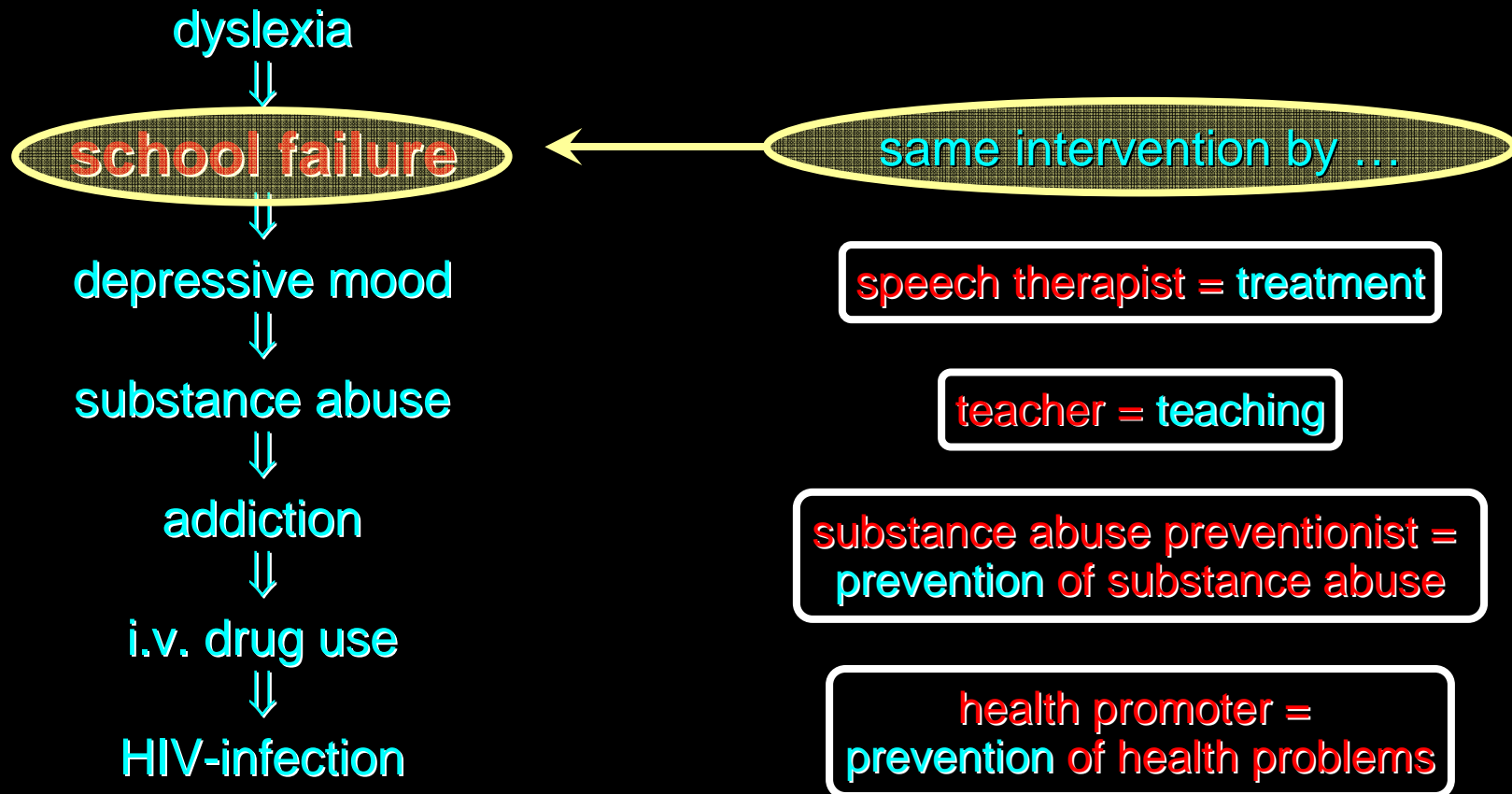


Mrazek, P.J.; Haggerty, R.J. (eds) (1994): Reducing Risks for Mental Disorders. Frontiers for Preventive Intervention Research - report of the Committee on Prevention of Mental Disorders Division of Biobehavioral Sciences and Mental Disorders Institute of Medicine. National Academy Press, Washington DC

# Is an action “treatment” or “prevention” ?



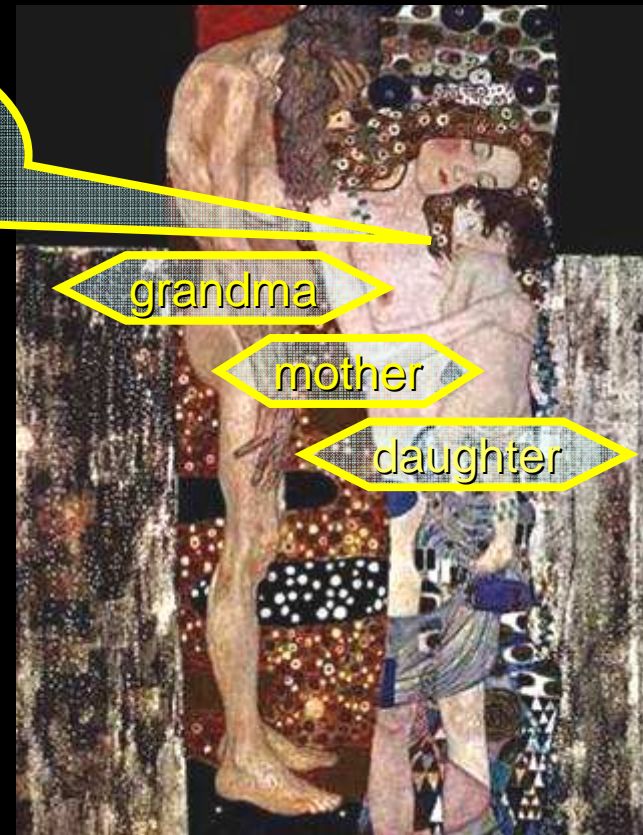
# How call an intervention ?



# Reification Fallacy

= hypostatisation of a relational term


Mom!  
Why do you call  
grandma „mother“ ?



Gustav Klimt: 1905 Three Ages of a Woman



## Semantic Confusion !

- 
- “prevention” as **specific activity**: Preventing what ?
  - “prevention” as **global activity**: What does it include ?
  - “prevention” as **certain profession**: Which professions ?

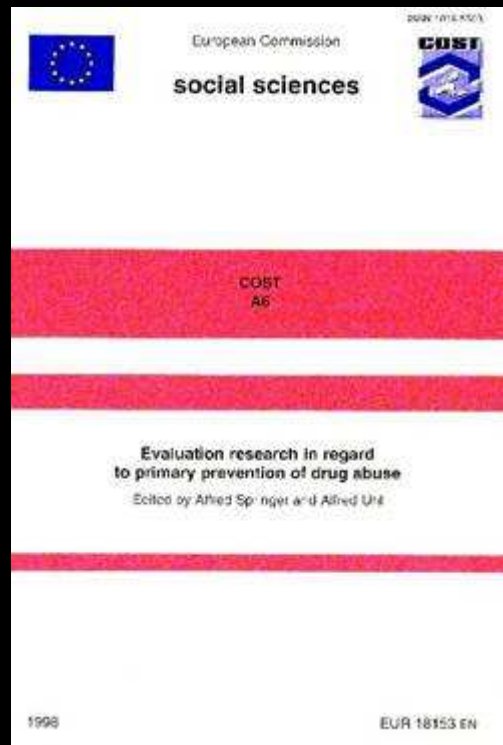
E.C. Escher 1960: Ascending & Descending



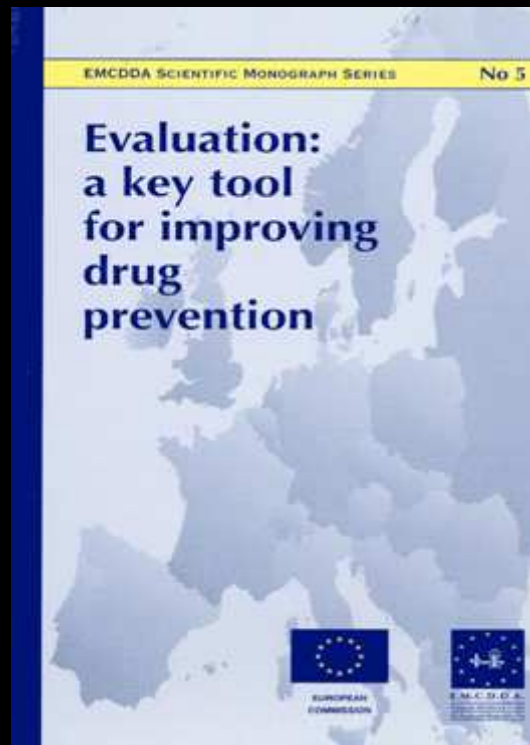
**(2)**

**What is  
„Evaluation“ ?**

# Evaluation Theory!



Uhl et al. (1998): Evaluation of primary Prevention in the Field of Illicit Drugs: Definitions - Concepts - Problems. COST A6, Brussels



Uhl (2000): The Limits of Evaluation , EMCCDA, Lisbon



German Evaluation Society (2003): Evaluation Standards. (DeGEval), Cologne



# Term “Evaluation”

- everyday context  
proving effectiveness of evaluand  
(= measure, technique, programme)  
(quite narrow – but mostly not feasible)
- professional context  
assessing anything to understand the evaluand better:  
(rather modest – and always feasible)

# Structure of a large Company

- Administration
- Research
- Production
- Advertisement
- Sales



Goofy as new Boss

# Structure of a large Company



Gyro Gearloose

- Production (administrative Processes)
- Production (of Research)
- Production (of Goods)
- Production (of Advertisement)
- Production (of Sales)

# Structure of a large Company

- Production
- Production
- Production
- Production
- Production

Beispiel: Autowerkstatt



Jerry Lewis

# Semantic Confusion !



„Tower of Babel“ Valckenborch (1595) in „old masters in Dresden“



# DTME - classification of Evaluation

(COST-A6-Workgroup)

- **Data dimension** (Type of data collection)
  - structural data
  - process data – **process evaluation**
  - expected outcome data – **outcome evaluation**
  - unexpected outcome data
  - context data
- **Time dimension** (State of evaluand)
  - concept phase = pre-formative phase
  - development phase = formative phase – **formative evaluation**
  - testing phase = first summative phase – **summative evaluation**
  - routine phase = second summative phase
- **Methodological dimension**
  - description
  - exploration
  - hypotheses testing
- **Evaluator dimension**
  - internal evaluation**
  - external evaluation**



# Wrong Conceptions about Evaluation

(Evaluation Standards – German Evaluation Society )

- **Always plan evaluation prior to project start**  
(commonly not feasible)
- **Always have precisely predefined goals**  
(commonly neither sensible nor feasible)
- **Every evaluation has to prove effectiveness**  
(usually neither sensible nor feasible)
- **The evaluator has to judge the evaluand**  
(sometimes explicitly ruled out – sometimes desired)
- **Only standardized methods are legitimate**  
(systematic, data-based and competent – but not standardized)



# Wrong Conceptions about Evaluation

(Evaluation Standards – German Evaluation Society )

- **Always quantitative statistics and significance tests**  
(understandable and criticisable – qualitative is accepted as well)
- **Always for specific practical application**  
(also gaining knowledge without intention to apply it)
- **Always aim at emancipation and not control**  
(not necessarily – but nice if feasible)
- **Always guarantee transparency for all participants**  
(not necessarily – but nice if feasible)



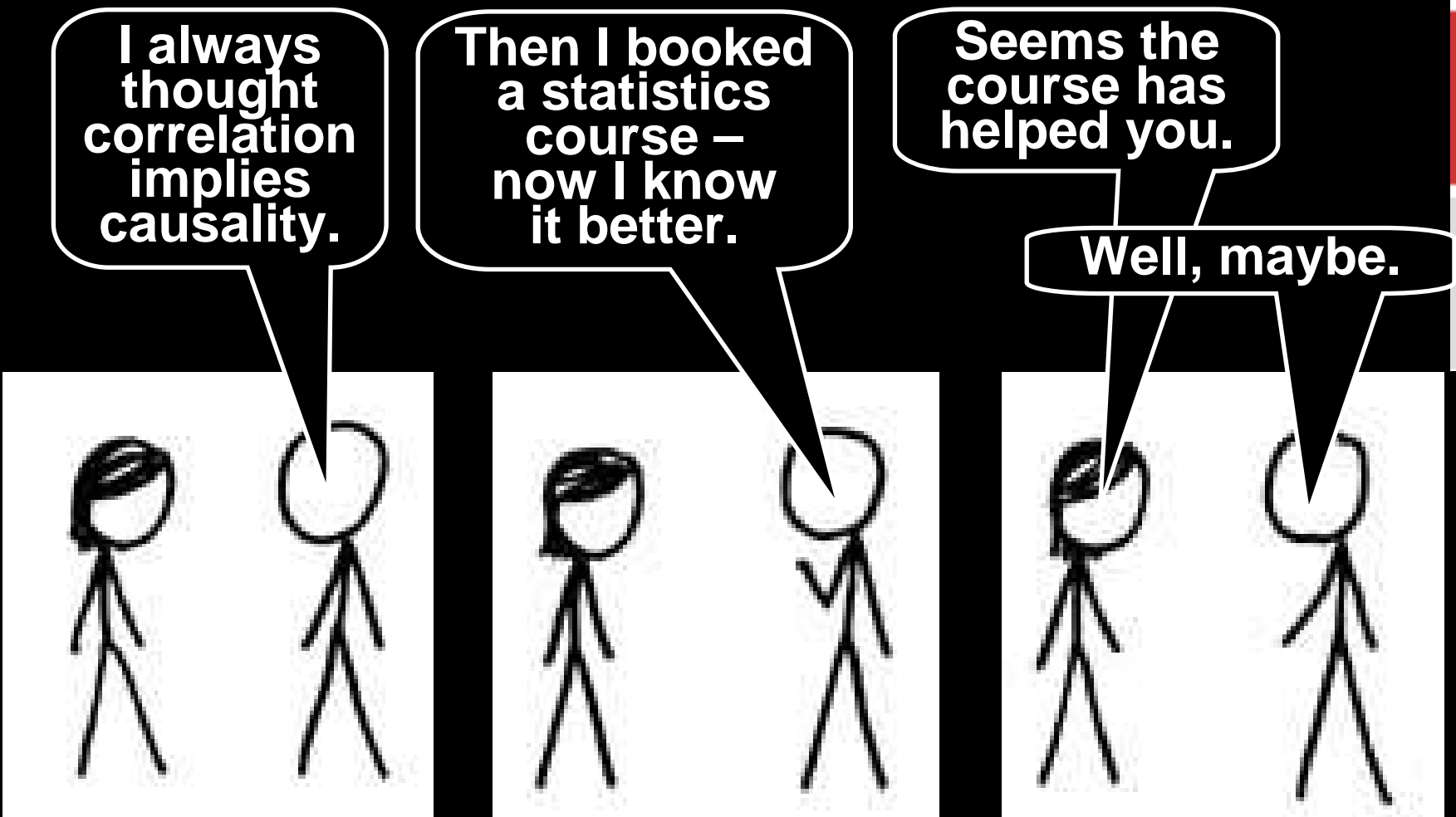
„Research methodology resembles grammar:  
Errors everybody makes,  
turn into new rules ...“

(feely based on André Malraux – about Politics)



**(3)**

## **Research Limitations**



I always  
thought  
correlation  
implies  
causality.

Then I booked  
a statistics  
course –  
now I know  
it better.

Seems the  
course has  
helped you.

Well, maybe.

Anything is possible if you don't know what you're talking about !  
Green's Law of Debate



# Competing Influences

Well controlled experiments are hardly feasible !

In an open system,  
there is no control of other variables,  
thus what we can know is quite limited.



# Complexity

Effects are a result of a **complex interaction**  
of various influences !



# Generativity

Commonly no homogeneous but chaotic effects !

Non-linear Theories,  
Chaos Theory, Catastrophe Theory, Complexity Theory



# The Power Dilemma

Enormous sample sizes needed.  
Non-significance does not mean no effect.

1% total lifetime prevalence over 10 years  
0.1% incidence per year  
20% relative effect, i.e. 1 less out of 2 times 5000  
360 000 minimum sample size

**provided:**  
**random sample not cluster sample (classes)**  
**no measurement error**

Uhl (2002): Schutzfaktoren und Risikofaktoren in der Suchtprophylaxe



# Statistical Regression

2/3 artificial spontaneous remission  
of alcoholism in one years  
due to imperfect reliability  
if there is actually no effect at all !

Uhl (2002): Schutzfaktoren und Risikofaktoren in der Suchtprophylaxe

# Long-term Nature of Effects

Relevant are long-term effects.  
Surrogate variables (proxies) are often problematic

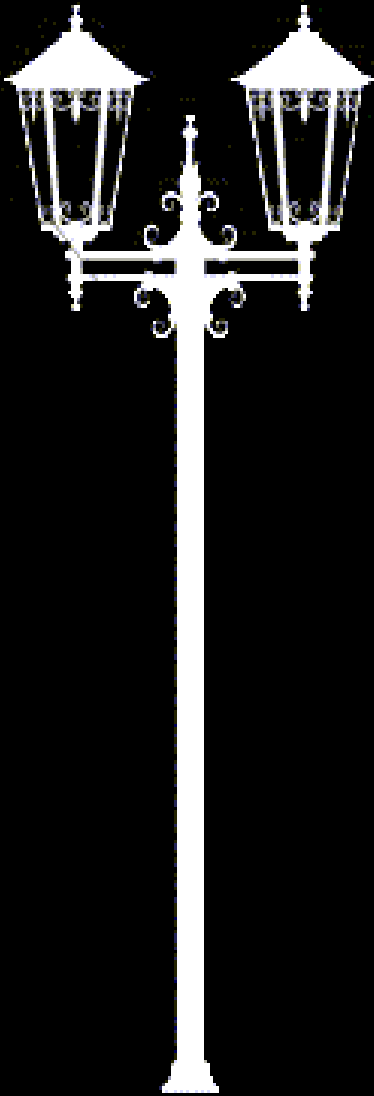


e.g. parrot effect:  
It is naïve to expect, that a parrot,  
trained to say „I hate bananas“,  
will not touch bananas in the future.“

e.g. Klingemann & Carter Sobell (2007):  
Promoting Self-Change from Addictive Behaviors. Springer, New York  
Uhl & Kűfner (2010) in preparation



# Assessing irrelevant aspects



We commonly do not measure  
what is relevant,  
(„soft skills“ = experiences, skills)  
but things that are easy to assess!  
Commonly just anything

e.g. Input = name of programme or technique  
Output = contentment / questionable proxies





**(4)**

## **Sensible Strategies**

Some things are too complex  
to be assessed with  
quantitative methods !!!  
„Soft Skills“



If we fail to do all things,  
when we don't know the reason for  
them or if we are unable to justify  
them, we would probably soon be  
dead.



Friedrich Hayek (1996): Die verhängnisvolle Anmaßung



Most of our knowledge  
is acquired by stepwise learning  
(hermeneutic phenomenological approach)

The hermeneutic process  
– respectively circle –  
contains a paradox:  
Whatever we want to understand must  
somehow be understood before.

We engage in a 'hermeneutic spiral'.



Car doesn't move → Tank empty



important:  
it is no prove!

- \* Fantasy
- \* Pre-knowledge
- \* Assumptions



ignition plug

logically speaking a fallacy!

## In a Diabetes Clinic

It makes **no sense**  
to prove that insulin works !

We have to check,

- if **theory** is applied correctly
- if **hygiene standards** are met
- If waiting times are **acceptable**
- If the staff **informs patients** adequately
- etc.






# Shelters for Abused Women



It makes **no sense** to prove that domestic violence is reduced !

We have to check,

- if women **know** about the facility
- if the **capacities** are sufficient
- If **security** is guaranteed
- etc.




The public prefers to listen to **scientists** who give **confident answers** to questions and make confident predictions ... So ... the experts who talk publicly about politically contentious questions tend to speak more clearly than they think.

They make **confident predictions** about the future, and end up **believing their own predictions**. Their predictions become **dogmas** which they do not question. ... That is why **heretics** who question the dogmas are needed.



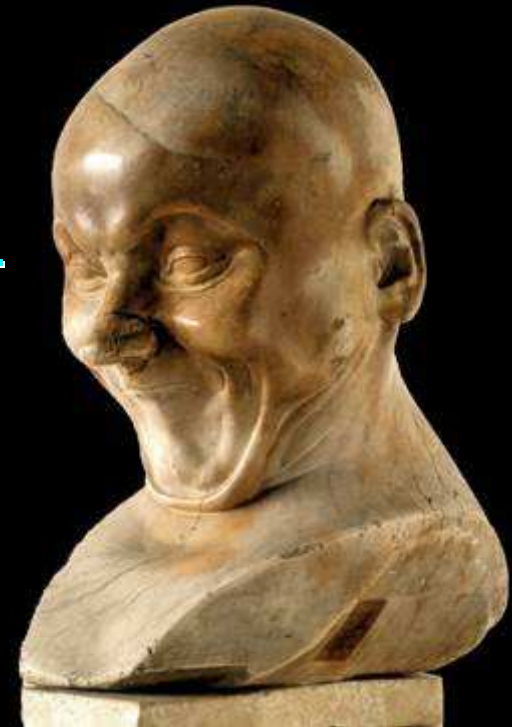
Freeman Dyson (2007) Many Colored Glass:, University of Virginia Press, 2007



In contrast to “honest stupidity” there is a form of “higher stupidity”. The latter does not result from lacking intelligence but in the contrary is a sign of intelligence and at the same time signals its failure, assuming an attainment, that is not existent.

Robert Musil (1937) über die Dummheit

Experts are commonly individuals who provide information so frequently that they have no time to collect information.





**(5)**

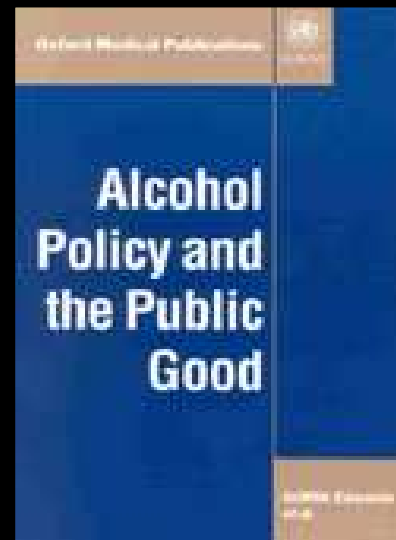
## **Evidence Based Alcohol Policy or Restoration of Paternalism**

# Alcohol Control Policy

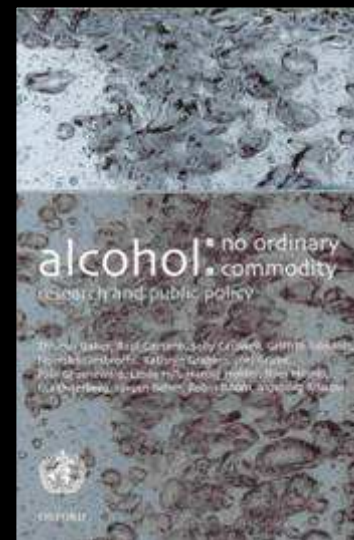
## 4 Main Publication



1975



1995



2003



2006

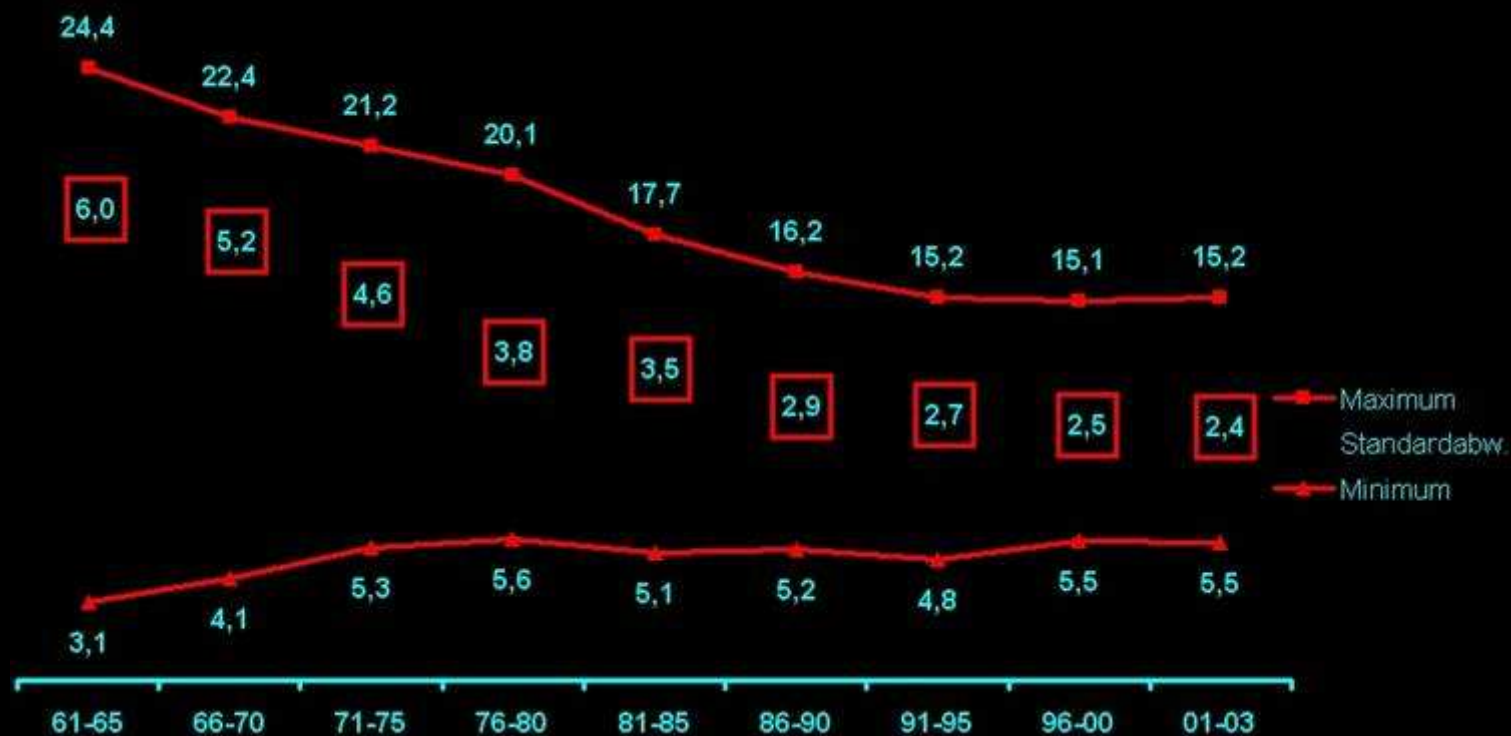
publications coining the positions of WHO & EU



## The books offer:

1. comprehensive literature views (+)
2. critical and well-founded analyses (+)
3. ... but over-simplifying conclusions supporting advocacy for a Northern European perspective (-)

# e.g. European Convergence





# (6)

## Evalopathy



# Evalopathy

... is the pathological drive to evaluate  
even if where it is **not possible**  
and  
**where it makes no sense;**  
  
**with inadequate methods**  
and  
**irrational goals**

Uhl, A. (2000): Evaluation vs. Evalopathy:  
Support for Practical Improvement vs. Irrational Nuisance



**Evalopathy turns into a new form of bureaucracy.  
The required documentation of output  
prevents increasingly that these  
accomplishments can be reached.**



## performative Self-immunisation

Certain words are **magic words** blocking our thinking, making it virtually impossible to grasp what they actually stand for.


This works based on a simple scheme:

The words **do not denote** what the term suggests and **hide** what they actually indicate.

This peculiar delusion works only, due to the principle of “**Performative Self-immunisation**”.

Persons using these terms have inevitably won, since **rejecting the terms** is only possible at the **cost of self-destruction**. Nobody can afford to be suspected of being against measuring effectiveness, improving quality, considering cost-effectiveness – and related to evidence-based policies – to be against using the best available evidence.

Liessmann, K.P. (2009): Theorie der Unbildung Die Irrtümer der Wissensgesellschaft (2. Auflage). Piper, München

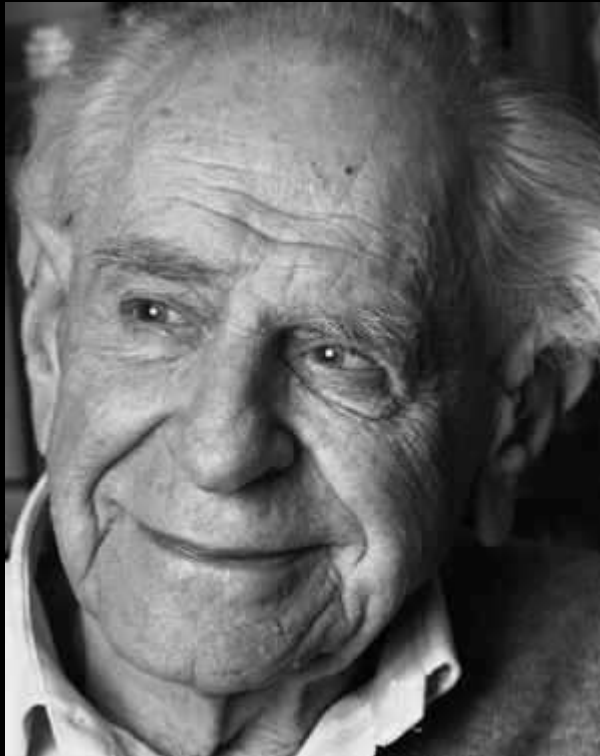


The most important result from brain research:  
„We have a brain and ...



... we should use it !”

## cross-thinking – scepticism – self-critics



Rationality can be explained  
as openness for criticism  
- as readiness to allow being  
criticised and as  
urge to criticise oneself  
Karl Popper Ausgangspunkte

**critical rationalism**  
**Karl Popper**

It is difficult to get a man to understand something when his salary depends on his not understanding it.

Upton Sinclair



**Stock Broker (2009)**



## What is needed

- theory and experience
- basic methodology and reasoning
- awareness of implicit ethical decisions
- divergent thinking – many possible explanations (abduktive thinking)
- frustration tolerance and endurance to try different approaches over and over
- ambiguity tolerance and courage to admit uncertainties (honesty)



## What is needed

### Think tanks – centers of excellence

- research programs = long-term perspective
- time to discuss, learn and disseminate
- time to train new-comers
- basic funding (plus project funding)
- more courage to challenge others and fear that bad reasoning does not pay



Wisdom is perceiving all things to be important  
but nothing totally serious.

Arthur Schnitzler



Thank you for your attention