

# The analysis of a set of questionnaires composed of 151 items on salt-, sweet- and fat-liking using LISREL-type and PLS-PM techniques leads to equivalent results

Amélie Deglaire<sup>1</sup>, Pascal Schlich<sup>1</sup>, Caroline Méjean<sup>2</sup>, Serge Hercberg<sup>2</sup>

26<sup>th</sup> of July, 10<sup>th</sup> Sensometrics, Rotterdam

<sup>1</sup>Food and Behaviour Research Center, UMR1324 INRA, UMR6265 CNRS, Université de Bourgogne, Agrosup, Dijon, France

<sup>2</sup> Nutritional Epidemiology Research Unit, Bobigny, France



- ❑ **Fat, sugar or salt** overconsumption: risk for health (WHO/FAO, 2003)
- ❑ Dietary consumption and **hedonism**
  - Positive correlation between liking for fatty foods and dietary fat intake (Drewnowski et al, 2000; Geiselman et al, 1998; Ledikwe et al, 2007; Raynor et al, 2004)
- ❑ **Measure** related to saltiness, sweetness and fattiness
  - Consumption : standardized tools (nutritional epidemiology)
  - Overall liking : no standardized tool (sensory science)
- ❑ **EpiPref Project** (funded by the French National Agency of Research)
  - develop and validate tools (sensory tests and a questionnaire) to measure the **overall liking** towards saltiness, sweetness and fattiness

- **Development and validation of a set of questionnaires** measuring the **overall liking** towards **saltiness, sweetness and fattiness (PrefQuest)**
  - *First stage: pilot study* (198 participants , 2009)
    - Feasability and reproducibility demonstrated, internal validity explored
  - *Second stage: application to thousands* of French people (n >40 000, 2010)
    - Analysis of this dataset
  
- **Structural Equation Modeling (SEM)**
  - LISREL-type (Linear Structural Relationship, covariance-based SEM)
  - PLS-PM (Partial Least Squares Path Modeling, component-based SEM)
  - Comparison of the latent variable score computation between these two SEM (Tenenhaus et al, 2005)
  - To our knowledge, never undertaken with such a high number of observations and such complex models

## **Analysis of the PrefQuest administered to a large population sample (n= 43 807)**

- ❑ **Exploration of the questionnaire structure**
  - identify the latent factors underlying the questionnaire items for one sensory sensation (sweet, salty, fatty-sweet, fatty-salty)
- ❑ **Validation of this structure**
  - statistically confirm the relationship between the observed and latent variables and assess the interfactor correlations
- ❑ **Calculation of an individual score** on each sensation using both LISREL-type and PLS-PM techniques
- ❑ **Comparison** of the two techniques

## Questionnaire Design

**151 items scattered into**

- ❑ **Four sensory sensations**
  - Sweetness
  - Saltiness
  - Fattiness and sweetness
  - Fattiness and saltiness
  
- ❑ **Four types of questions for each sensation**
  - Liking for specific foods
  - Preferred level of seasoning
  - Preferred dishes in a menu
  - Dietary behavior questions

# QUESTIONNAIRE CONTENT

## 1. Foods

### □ Question

How much do you like...

Olives:

I do not like them at all         I like them very much (+)  I have never tasted

- 9-point scale

### □ Food selection

- Foods from various food families with a high content in:
  - Sugar (drinks, desserts, biscuits,...)
  - Fat and sugar (pastries, desserts, chocolate, ...)
  - Fat and salt (cured meats, sauce, cheese, ...)
- Pretests
  - foods tasted by 85-90% of the subjects
  - the most discriminant foods

Sensation	nb items
Sweet	20
Fatty-sweet	22
Salty	0
Fatty-salty	31

## 2. PREFERRED LEVEL OF SEASONING

### Question without picture

How do you prefer your steak...

- not salty at all
- not too much salty
- moderately salty
- quite salty
- very much salty
- I do not like steak

- 6- or 5- point scale

How do you prefer your strawberries...



Tick the box that corresponds to the way you prefer your strawberries:

with no whipped cream  0  1  2  3  4  5 with a lot of whipped cream

Sensation	Nb items
Sweet	9
Fatty-sweet	10
Salty	9
Fatty-salty	12

### 3. Menu

#### □ Question

**Select the dishes you prefer.** You can select up to 4 dishes into the list below.  
If you do not like anything in the list, tick the box labelled as « none »

- Beef steak (entrecôte) and its roquefort or béarnaise sauce
- Beef steak (entrecôte) cooked with thyme
- Chicken in mushroom sauce
- Chicken steamed with lemon
- Pork with creamy mustard sauce
- Pork on the grill
- White fish and its buttery sauce
- White fish in papillote with estragon
- None

- 4 dishes fatty-salty versus 4 dishes not fatty-salty
- *Score*: number of fatty and salty items / total number of items chosen

#### □ Types of dishes per sensation

- *Sweet*: drinks for appetizer (alcohol), dessert, cold drinks (no alcohol)
- *Fatty and Sweet*: dessert, snack dessert, hot drinks
- *Salty*: appetizers, meat
- *Fatty and salty*: meat, side dishes, italian food

Sensation	nb items
Sweet	3
Fatty-sweet	3
Salty	2
Fatty-salty	3



## 4. Dietary Behavior Questions

### □ Question

#### ▪ Behavior/habits

**Do you add some salt to your meal without tasting it?**

never   
  rarely   
  some-  
times   
  often   
  always

#### ▪ Context

**You buy a « butter-ham » sandwich in a bakery. Once in the street, you realize it's a sandwich with ham but no butter. Do you mind eating this sandwich without butter?**

not at all                                        very much

#### ▪ 5- or 9-point scale

Sensation	nb items
Sweet	4
Fatty-sweet	4
Salty	6
Fatty-salty	3

## 5. OVERVIEW

### Number of items per sensory sensation and type of question

Sensation	Foods	Preferred level of seasoning		Menu	Behavior questions	Total
		with pictures	without pictures			
Sweet	20	8	1	3	4	36
Fatty-sweet	22	9	1	3	4	39
Salty	0	7	2	2	6	17
Fatty-salty	31	10	2	3	3	49
<b>Total</b>	<b>73</b>	<b>34</b>	<b>6</b>	<b>11</b>	<b>17</b>	<b>141</b>

Preferred level of seasoning : 10 similar questions were included with or without pictures in order to study the picture effect

- Overall, 141 unique items scattered into 4 sensory sensations

- ❑ **Internet-based** through the national survey of Nutrinet-Santé (*Nutritional Epidemiology Research Unit*)
- ❑ **March 23<sup>rd</sup> – May 10<sup>th</sup> 2010:** 43 807 participants to the Questionnaire



<b>Age years</b>	<b>All participants % column</b>	<b>Women % line</b>	<b>Men</b>
<b>18-29</b>	19	88	13
<b>30-39</b>	21	82	18
<b>40-49</b>	19	80	20
<b>50-59</b>	20	78	22
<b>≥60</b>	20	58	43
<b>All participants</b>	100	<b>77</b>	<b>23</b>

## □ Data preparation and screening

- transformation into values within [0;1]
- screening for univariate normality (skewness<2, kurtosis<7; Curran et al, 1996; Kline, 2005) : deletion of 5 items
- dataset without any missing data : **n=8 800\***
  - **Cross-validation:** training base (n=4275) and test base (n=4285)

## □ Structure exploration

- Exploratory Factor Analysis (EFA) on the training base
  - SAS® proc FACTOR
  - maximum likelihood extraction
  - oblique rotation (Promax)

## □ Structure validation

- Confirmatory Factor Analysis (CFA) on the test base
  - SAS® Proc **CALIS** (Covariance Analysis of Linear Structural Equations)
  - **maximum likelihood** parameter estimation on the covariance matrix
  - second-order factor analysis (hierarchical model)

\* Missing data due to the « never tasted » point in the scale

## □ Score computation

- **Regression coefficients** computed on the test dataset
  - **CALIS** (SAS<sup>®</sup> software) - ML
  - **PLS-PM** (XLStat software) - reflective mode (A)
- **Latent variable score** for each subject (n=43 806\*) and sensory sensation
- **Missing data treatment**
  - replacement of the subject's missing value by the mean of the factor to which it belongs to

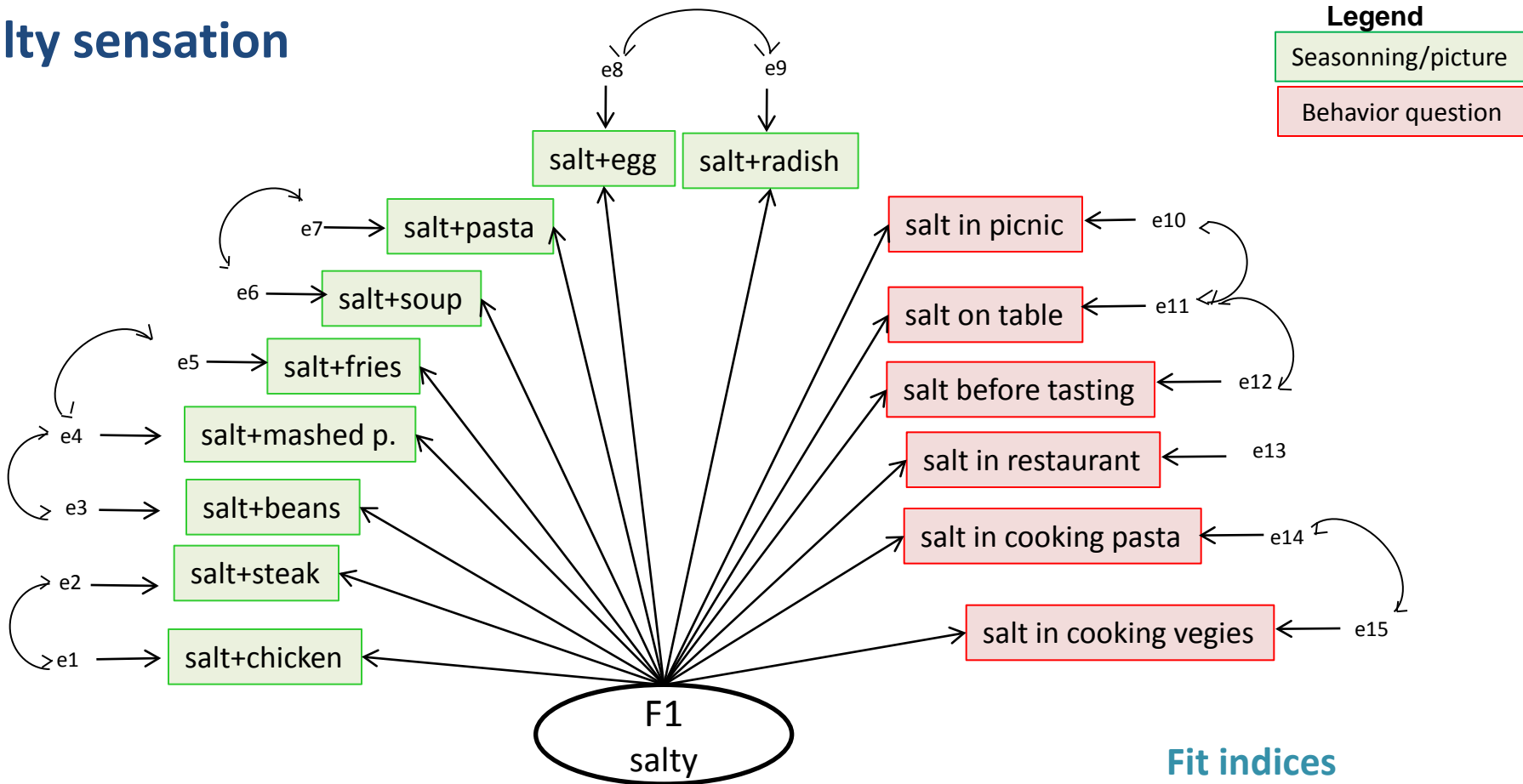
## □ Comparison between CALIS and PLSPM

- Interfactor correlations, regression coefficients and scores
- Scatter Plot and Pearson's correlation coefficients

*\* 1 observation discarded due to the high number of missing data*

# RESULTS/STRUCTURE OF THE QUESTIONNAIRE

## Salty sensation



- 15 out of 17 items selected (EFA, n=4275)  
Cronbach  $\alpha$ : 0.89
- Error covariances (order effect + question type)
- Model validated (CFA, n=4285)

Root Mean Square Error of Approximation (RMSEA)	0.06
Goodness of Fit Index (GFI)	0.96
Comparative Fit Index (CFI)	0.95
Non-Normed Fit Index (NNFI)	0.94

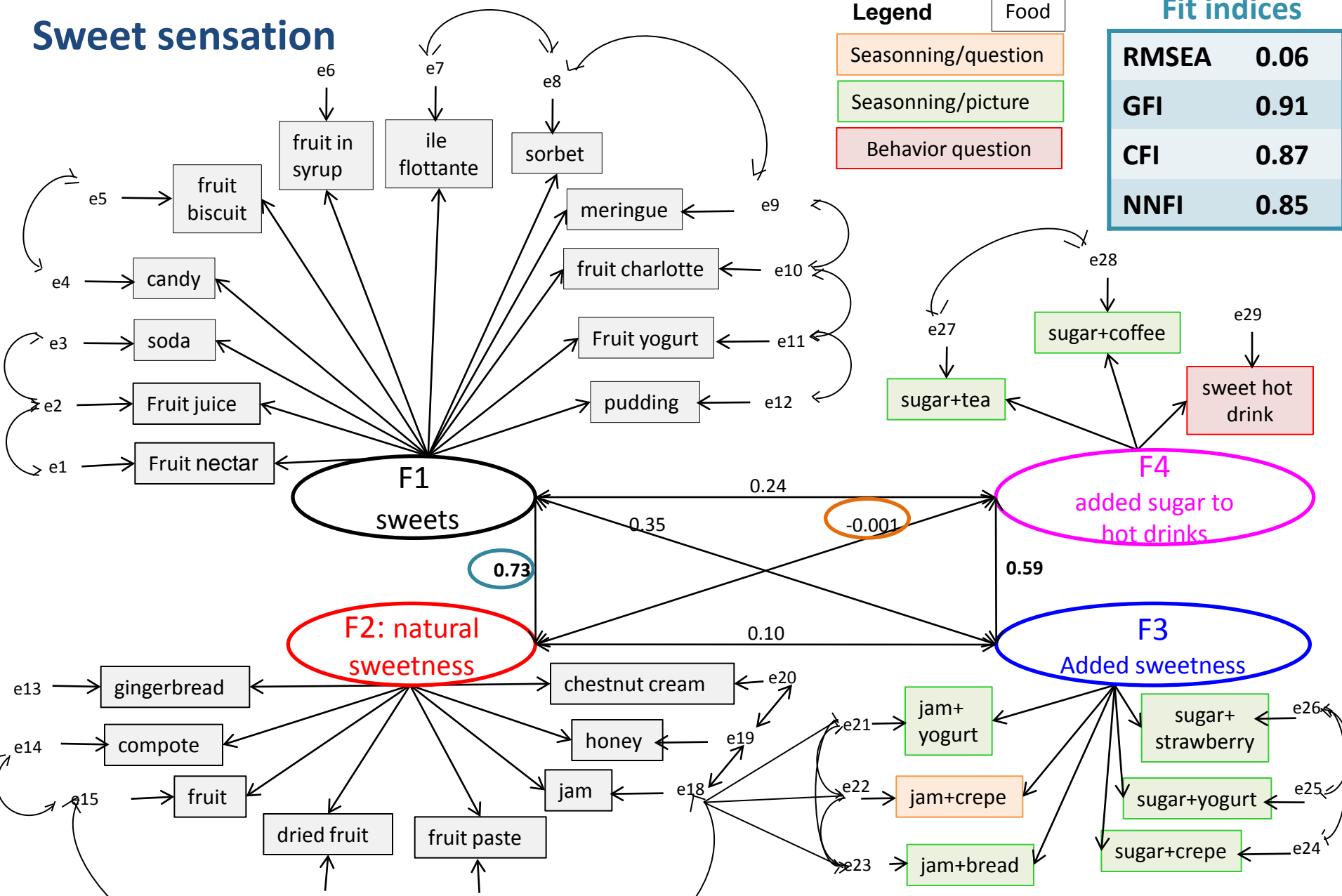
# Sweet sensation

**Legend**

- Seasoning/question
- Seasoning/picture
- Behavior question

**Fit indices**

RMSEA	0.06
GFI	0.91
CFI	0.87
NNFI	0.85

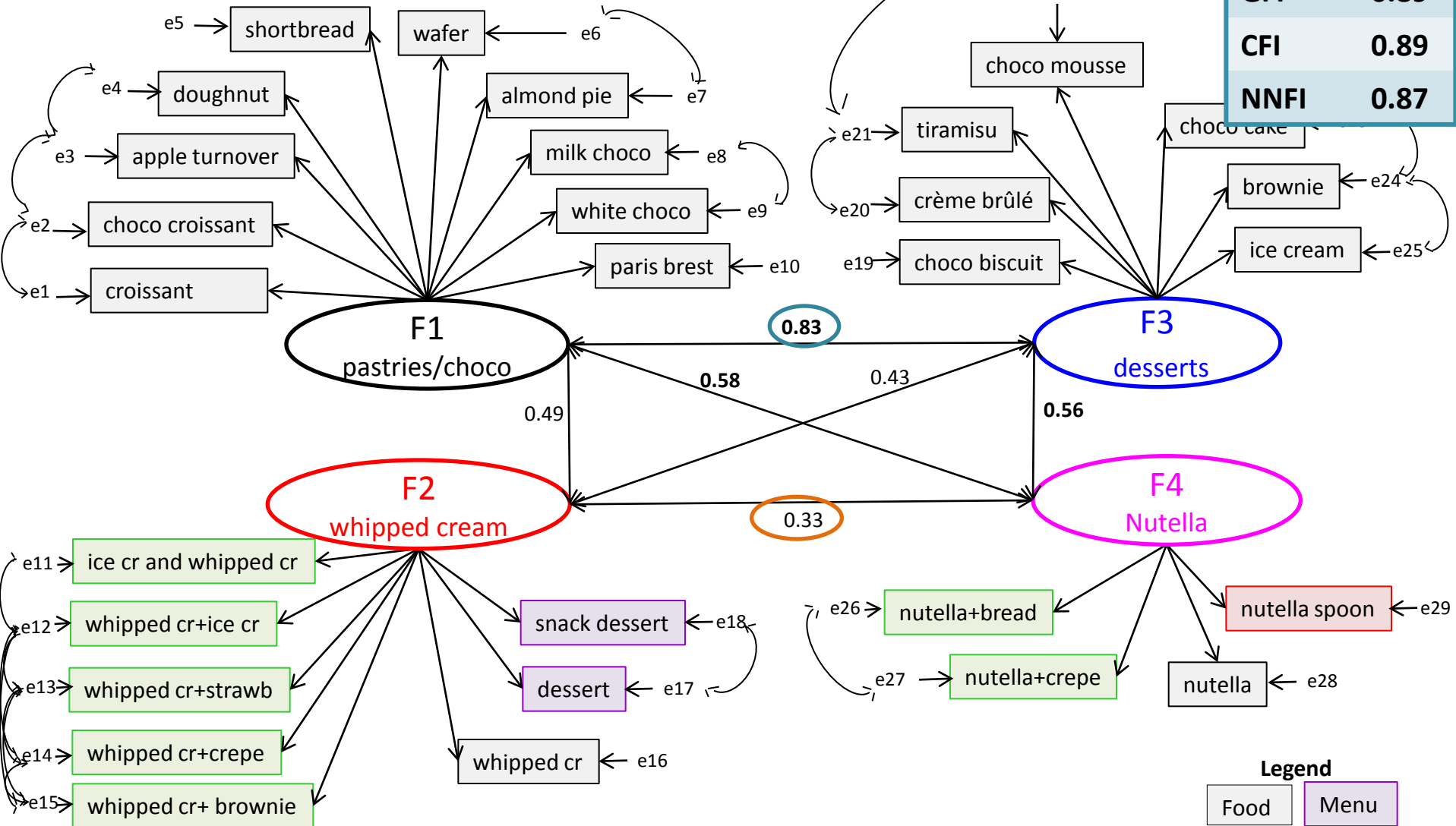


Pool of 36 items (CFA) selected (5/14 for F1, 2/7 for F2, 7/14 for F3, 2/7 for F4) for final questionnaire. Cronbach's alpha: 0.74 (F3)-0.85 (F1)

# Fatty and sweet sensation: first-order factors

## Fit indices

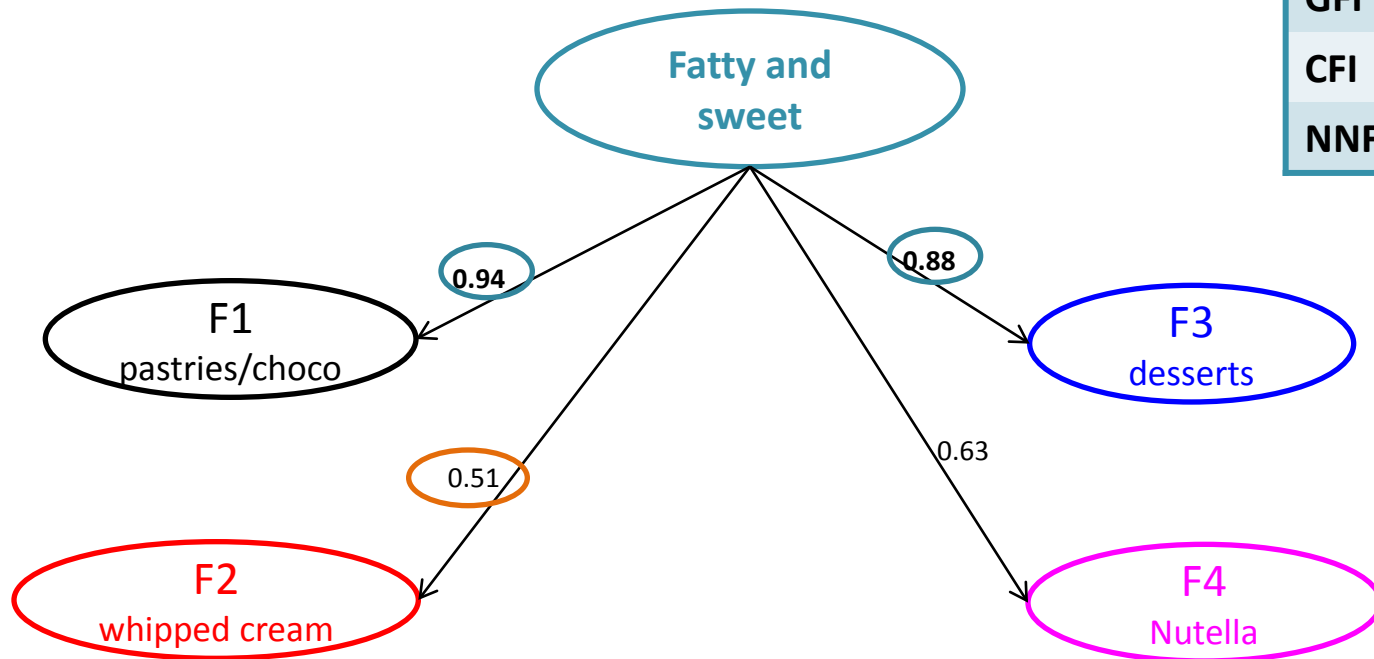
<b>RMSEA</b>	<b>0.07</b>
<b>GFI</b>	<b>0.89</b>
<b>CFI</b>	<b>0.89</b>
<b>NNFI</b>	<b>0.87</b>



Model validated (CFI = 0.85) for first-order factors (86 (F1)



## Fatty and sweet sensation: second-order factor



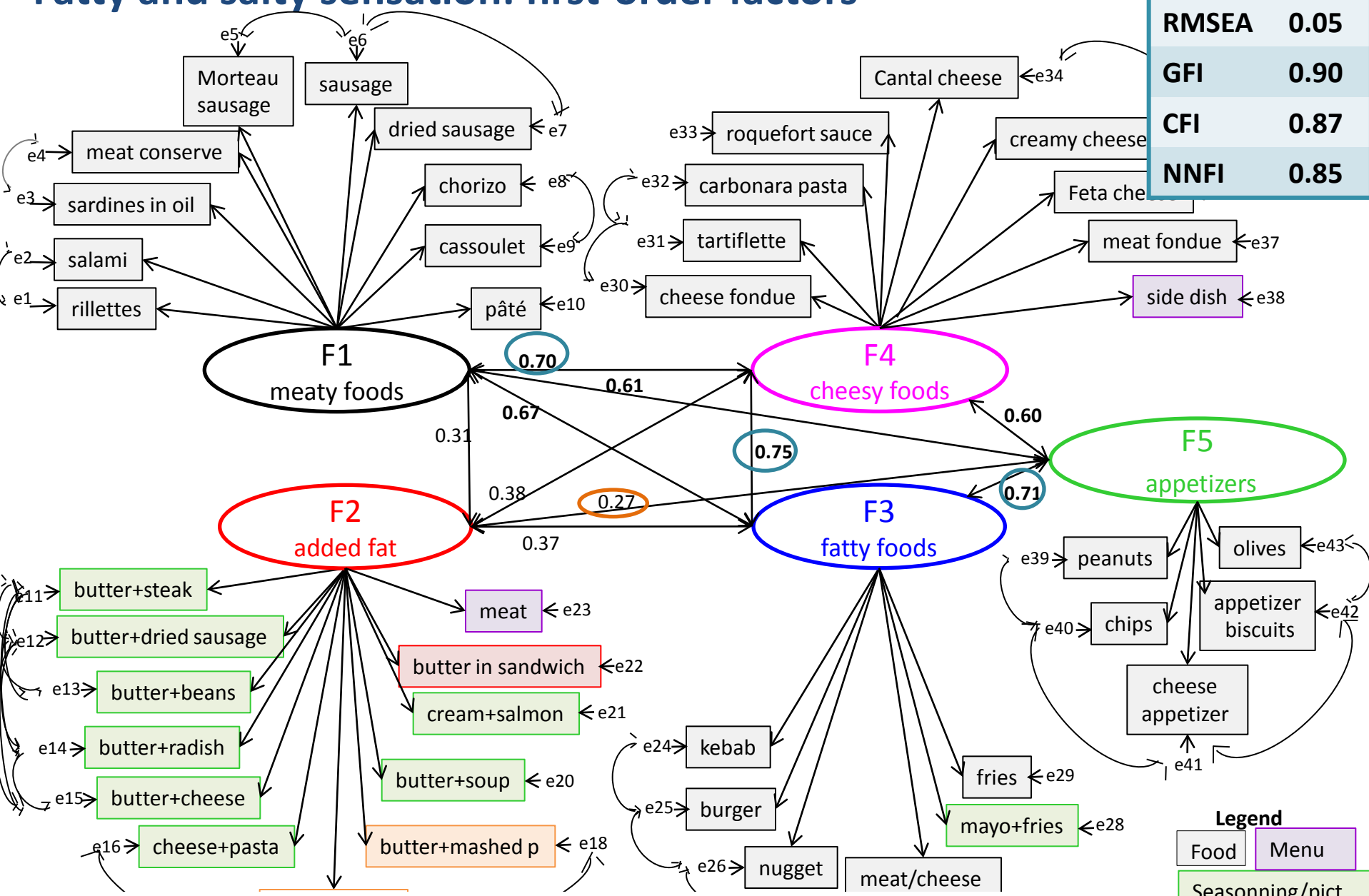
Fit indices	
RMSEA	0.07
GFI	0.89
CFI	0.89
NNFI	0.87

- Model validated for the second-order factor  
Cronbach  $\alpha$ : 0.91 (F5)

# Fatty and salty sensation: first-order factors

Fit indices

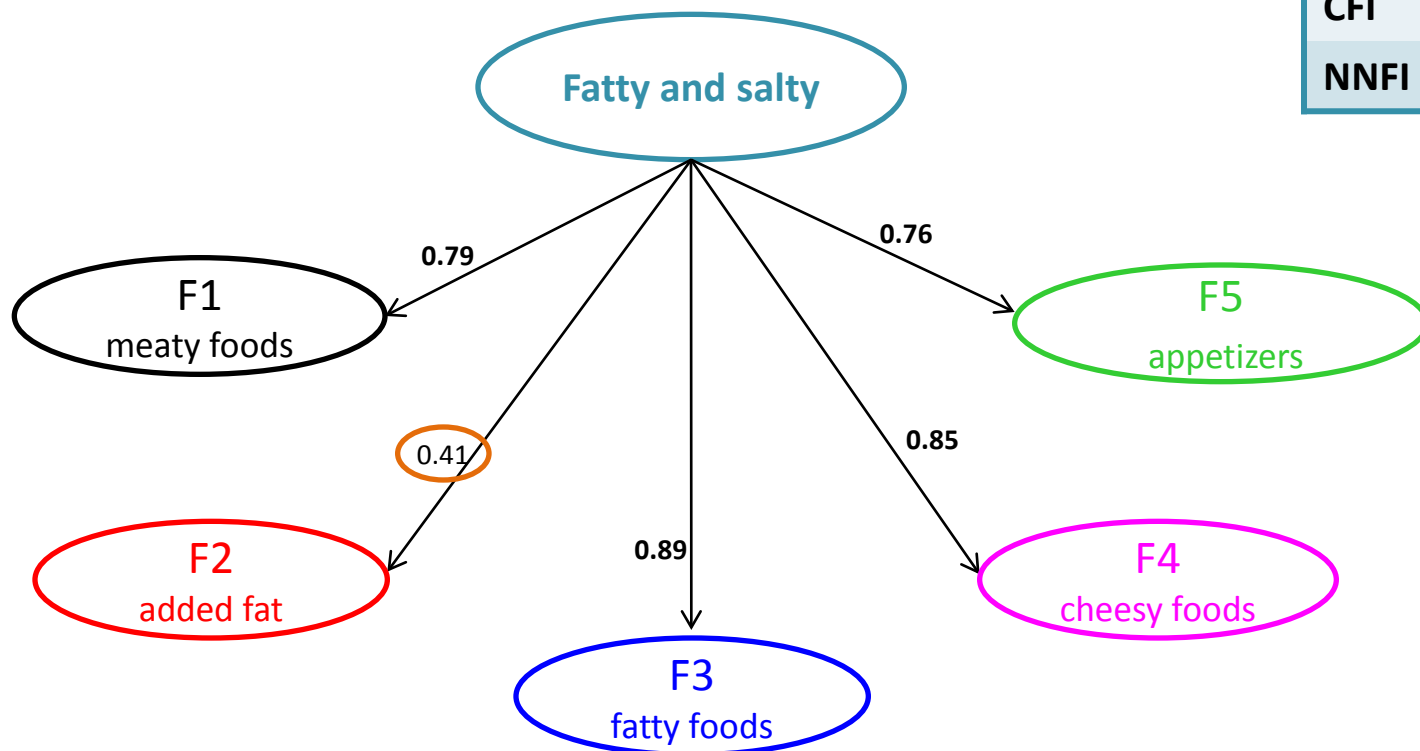
RMSEA	0.05
GFI	0.90
CFI	0.87
NNFI	0.85



Model validated (CFA, n=4285) for first-order factors

1)

## Fatty and salty sensation: second-order factor



### Fit indices

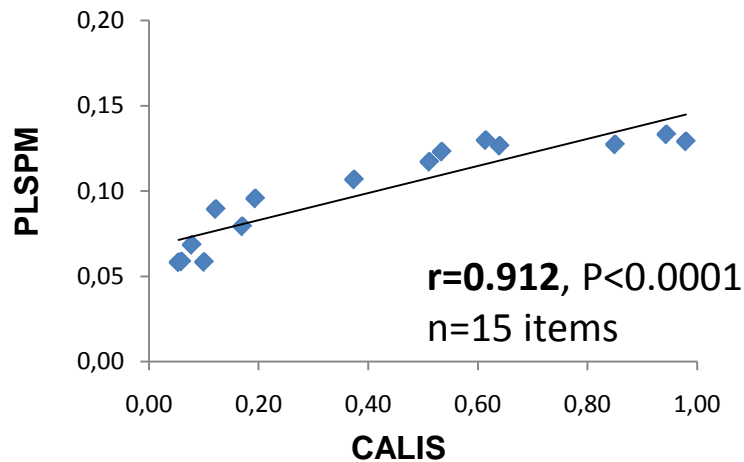
RMSEA	0.05
GFI	0.90
CFI	0.87
NNFI	0.85

- Model validated for the second-order factor  
Cronbach  $\alpha$ : 0.91 (F6)

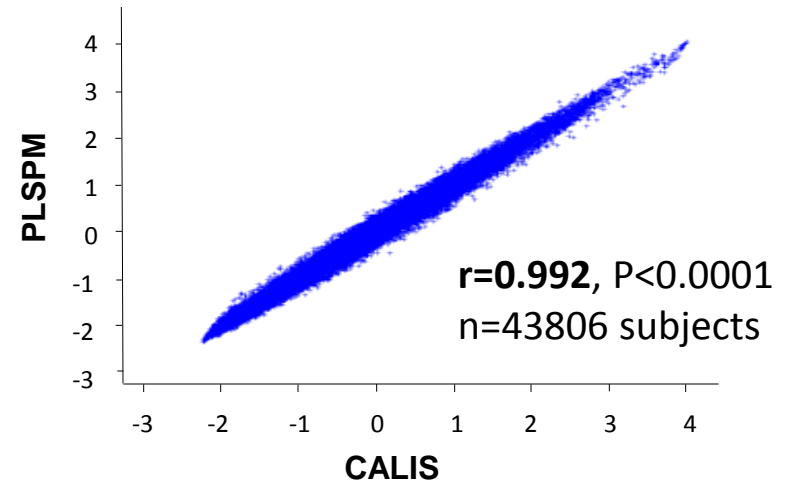
## Salty sensation: 1<sup>st</sup>-order factor

Correlation coefficients (Pearson,  $r$ ),  $P$ -value ( $P$ ), number of items/subjects ( $n$ )

### Regression coefficients



### Scores



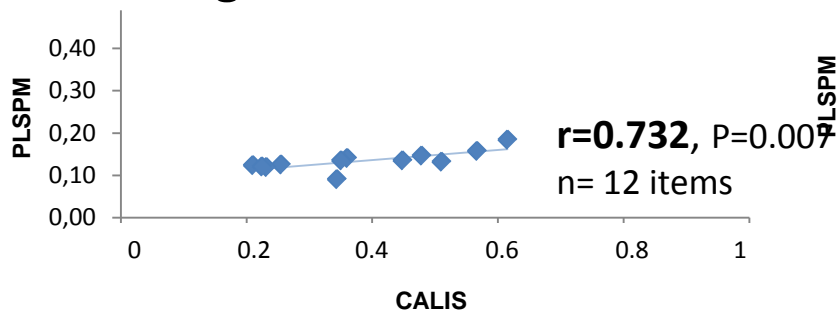
- High correlations for the regression coefficients and the scores

# Sweet sensation: 1<sup>st</sup>-order factors

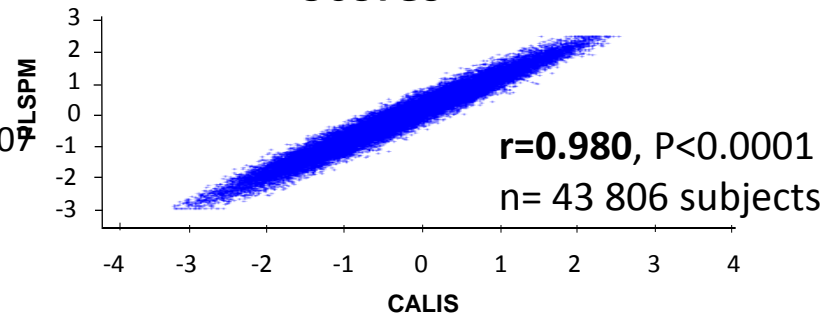
Correlation coefficients (Pearson,  $r$ ),  $P$ -value ( $P$ ), number of items/subjects ( $n$ )

F1  
Sweet  
foods

## Regression coefficients



## Scores



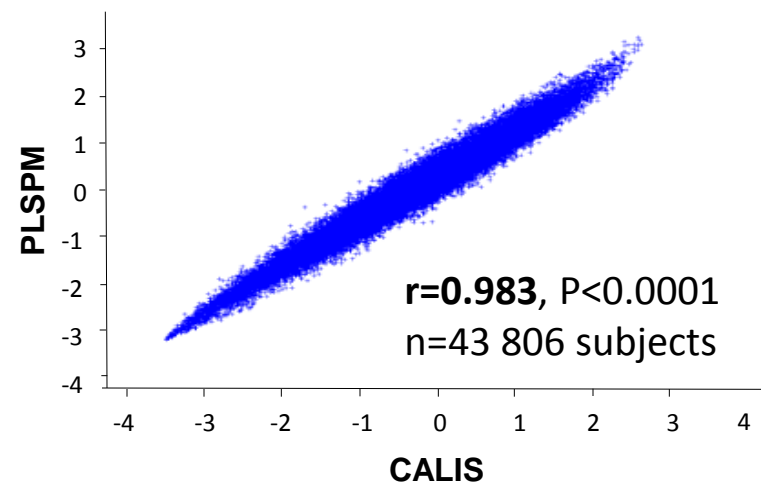
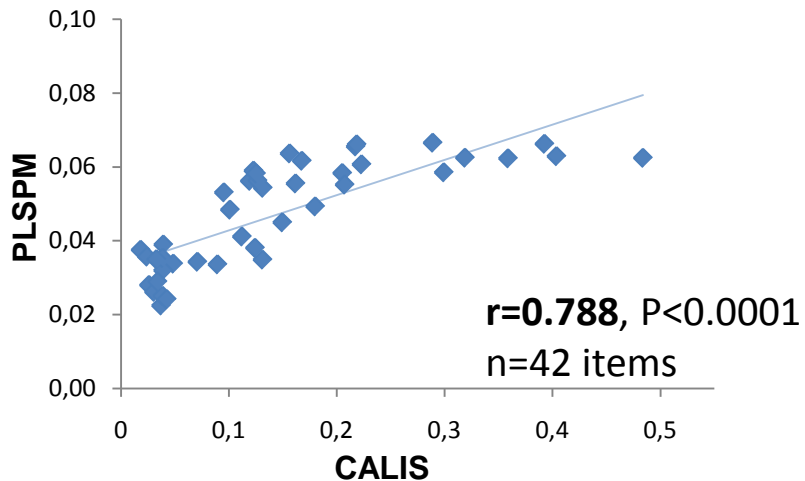
# Fatty and salty + fatty and sweet sensations: 2<sup>nd</sup>-order factors

Correlation coefficients (Pearson,  $r$ ),  $P$ -value ( $P$ ), number of items/subjects ( $n$ )

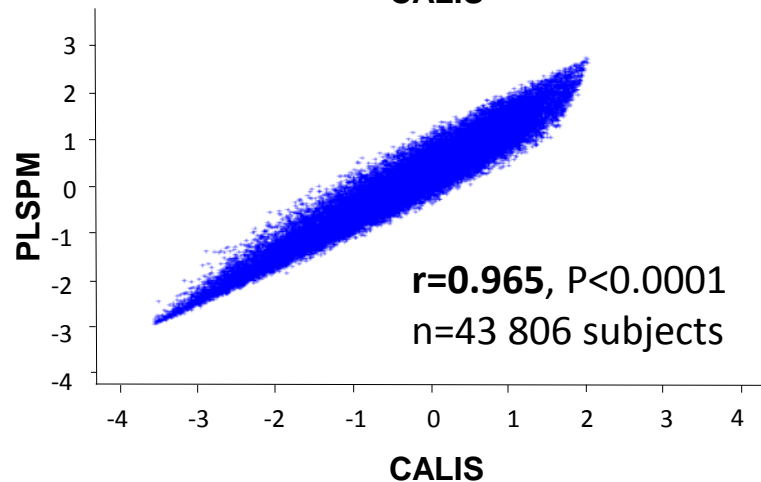
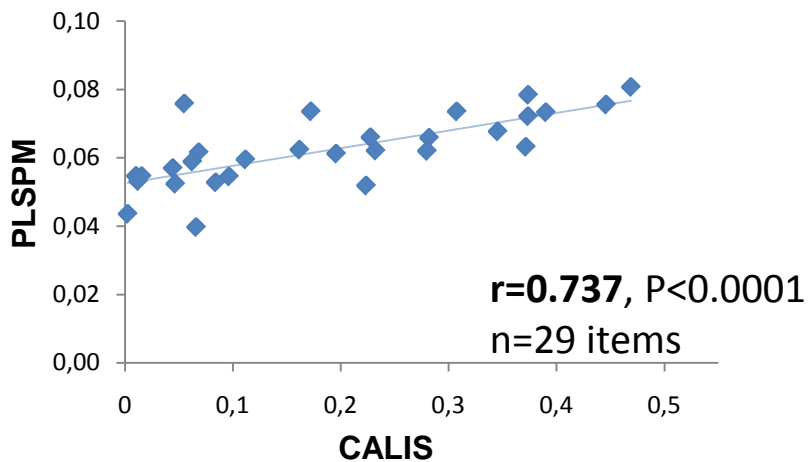
## Regression coefficients

## Scores

Fatty and Salty



Fatty and Sweet

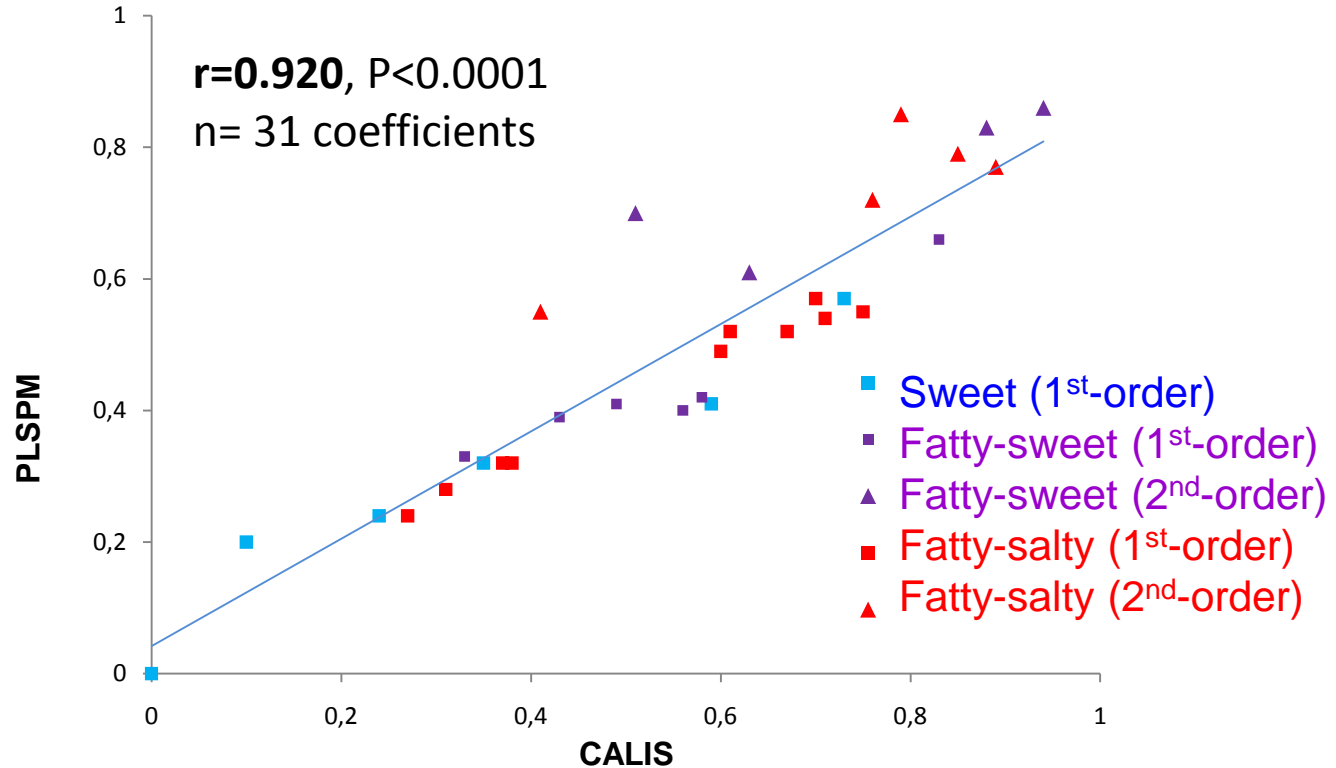


□ Better correlations for the scores than for the regression coefficients

# OVERALL COMPARISON BETWEEN CALIS AND PLS-PM

## InterFactor Correlations in the models for the sweet, fatty-sweet and fatty-salty sensations

Correlation coefficient (Pearson,  $r$ ),  $P$ -value ( $P$ ), number of interfactor correlation coefficients ( $n$ )



- ❑ 1<sup>st</sup> order factors: some underestimation in PLSPM compared with CALIS
- ❑ 2<sup>nd</sup> order factors: some overestimation in PLSPM
- ❑ In overall, InterFactor correlations similarly ranked in CALIS and PLSPM

# SCORE COMPUTATION TIME

Time required to complete the analysis and to calculate the individual scores

Sensation	1st-order factors	2nd-order factors	Nb items	Time computation (s) n=4 285		Time computation (s) n=43 806	
				CALIS	PLSPM	CALIS	PLSPM
Salt	1	0	15	0.2	13	4.0	101
Sweet	4	0	29	0.8	73	5.3	755
Fatty-Sweet	4	1	29	1.7	126	5.5	1320
Fatty-Salt	5	1	42	10.8	233	15.1	3758

- ❑ XLSTAT-PLSPM is about **60** (n=4285) up to **160** (n=43806) times longer than SAS<sup>®</sup> proc CALIS, but this is partly due to the bootstrapping
- ❑ **Multiplying the number of observations by 10** increases, in average, the computation time by **8** up to **11** for SAS<sup>®</sup> proc CALIS and XLSTAT-PLSPM.



# DISCUSSION & CONCLUSION

- CALIS and PLSPM leads to **similar interpretation in terms of individual scoring** over the 4 sensory sensations
  - Correlations in the same range as those in Tennenhaus et al. (2005)
  - Heterogeneity due to the **different process of score computation**
    - CALIS : linear combination of all the manifest variables in the model
    - PLSPM : linear combination of the manifest variables of the factor
- As known, **CALIS and PLSPM are complementary, with different objectives**
  - CALIS: model validation/better estimation of the structural model
  - PLSPM: score prediction/better estimation of the measurement model
  - Explain the lower degree of correlation for interfactor correlations and regression coefficients

- ❑ Computation of an overall score for the sweet sensation
- ❑ **Group effect** on the models
  - Gender/Age effect
- ❑ CALIS-ML: **data normality**
  - Univariate normality: data were moderately non-normal
    - Maximum of likelihood supposed to be robust (Curran et al., 1996)
  - What about **multivariate normality**?
    - Impact of multivariate non-normality on the ML estimation?
    - Rarely checked in the literature
    - Mardia multivariate kurtosis (Mardia, 1970): reliability with a high number of observations and manifest variables?

Thanks for your attention

