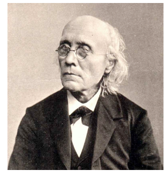
Estimating Perceptual Scales with Interval Properties

Kenneth Knoblauch[1] & Laurence T. Maloney[2]

1. Inserm, U846
   Stem Cell and Brain Research Institute
   Dept. Integrative Neurosciences
   Bron, France

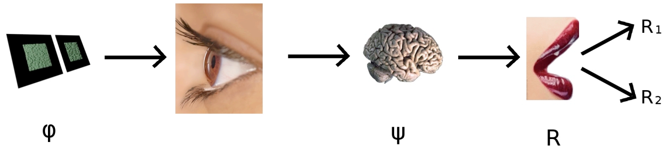
2. Department of Psychology
   Center for Neural Science
   New York University
   New York, NY 10003, USA

Psychophysics, qu'est-ce que c'est ?



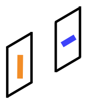
Gustav Fechner (1801 - 1887)

A body of techniques and analytic methods to study the relation between physical stimuli and the organism's (classification) behavior to infer internal states of the organism or their organization.



Classical Psychophysical Paradigm

Event — Observer — Response

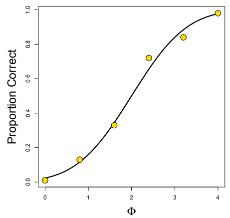


$\phi_i \in \{\phi_1, \cdots, \phi_n\}$

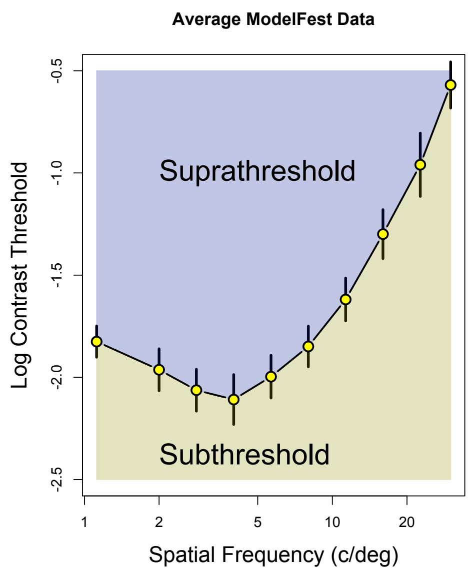
$\psi_j \in \{\psi_1, \cdots, \psi_m\} + \epsilon$

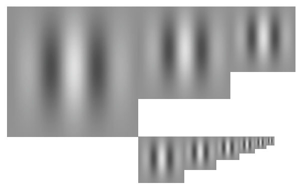
$\epsilon \sim \text{i.i.d.}$

$R_{ij} \in \{R_{11}, R_{12}, \ldots, R_{np}\}$



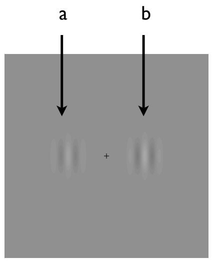
$Pr(R_{ij}) = f(\{\phi_i\}; \mathbf{\Psi})$, f is a Psychometric Function





http://vision.arc.nasa.gov/modelfest/

Paired Comparisons



Find numbers, (ψ_a, ψ_b), such that when b is judged of higher contrast than a, $\psi_b > \psi_a$.

Decision variable:

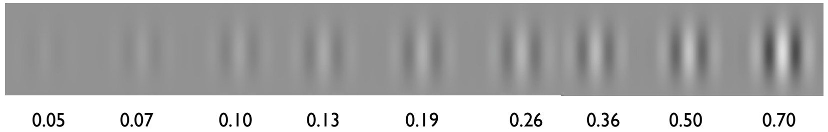
$$\Delta = \psi_b - \psi_a + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, \sigma^2)$$

Equal-variance, Gaussian Signal Detection Model

$$(\psi_1, \psi_2, \ldots, \psi_n)$$

Only yields an *ordinal scale*!
Any monotonic transformation of the above scale is equally valid!

Series of Suprathreshold contrasts



Two Questions to be considered in this talk

How to quantify the evolution of suprathreshold perception along a physical dimension (like contrast)?

How do different perceptual dimensions combine?

Methods that yield an *interval scale*

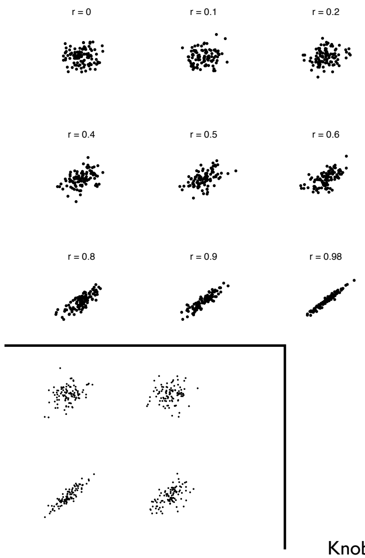
1) Difference Scaling: MLDS
(Maloney & Yang (2003). *J Vision*)

2) Conjoint Measurement: MLCM
(Luce & Tukey (1964) *J Math Psych*;
Ho, Landy & Maloney (2008) *Psych Science*)

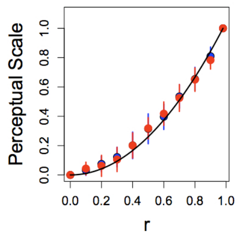
Both methods based on ordering intervals between stimuli and not on ordering stimuli, *per se*.

Ordering intervals leads to scales with interval properties, i.e, equal scale differences correspond to equal perceptual differences.

Difference Scaling: Correlation in scatterplots

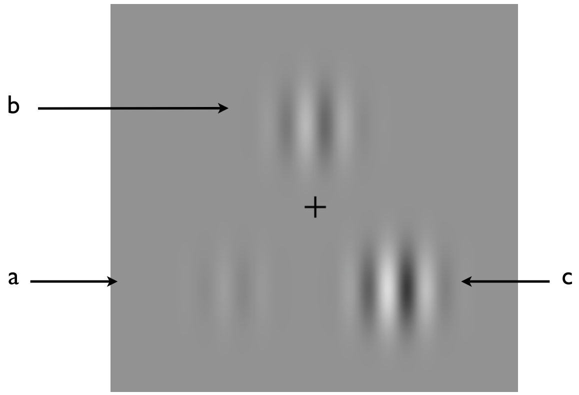


$\psi(r) \approx r^2$



Knoblauch & Maloney, *J. Stat. Soft.*, **25**, 1 - 26.

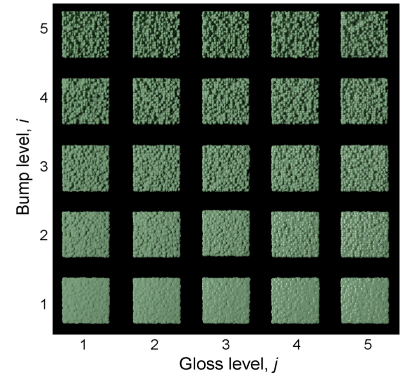
Difference Scaling: Triads



Between which pair, (a, b) or (b, c), is the contrast difference greatest?

Conjoint Measurement

Conjoint Measurement[1] is a psychophysical procedure used to estimate the interaction of perceptual scales for stimuli distributed along $n \geq 2$ physical continua.

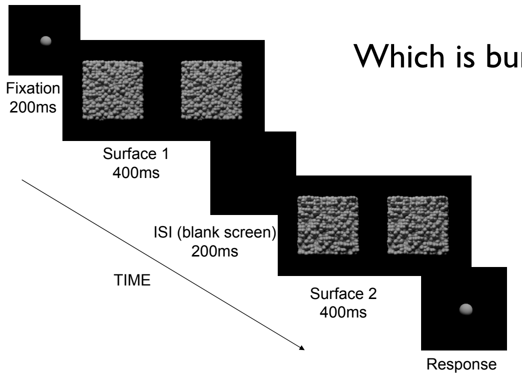


[1] *Luce & Tukey (1964) J Math Psych*

Ho, Landy & Maloney (2008) Psych Science

From a set of p stimuli varying along 2 dimensions, a random pair, (I_{ij}, I_{kl}), is chosen and presented to the observer as in this example.

The Task



Which is bumpier (glossier)?

Maximum Likelihood Difference Scaling: MLDS

The aim of the Maximum Likelihood Difference Scaling (MLDS) procedure is to estimate scale values, that best capture the observer's judgments of the perceptual difference between the stimuli in each pair.

The decision model

Given a quadruple, $\mathbf{q} = (a, b\,;\, c, d)$, from a single trial, we assume that the observer chooses the upper pair to be further apart when

$$\Delta(a, b\,; c, d) = (\psi_d - \psi_c) - (\psi_b - \psi_a) + \epsilon > 0$$

where ψ_i are estimated scale values, $\epsilon \sim \mathcal{N}(0, \sigma^2)$ and σ a scale factor.

Maximum Likelihood Conjoint Measurement: MLCM

The decision model

$b_1 \ g_1$ $b_2 \ g_2$ *Bumpier?*

$$\begin{aligned} B_1 &= \psi^b(b_1) + \chi^g(g_1) \\ B_2 &= \psi^b(b_2) + \chi^g(g_2) \\ \Delta &= B_1 - B_2 + \epsilon > 0 \Leftrightarrow \text{“First”} \end{aligned}$$

$$\epsilon \sim \mathcal{N}(0, \sigma^2)$$

Ho, Landy & Maloney (2008), Psych Science

Estimation of Scale Values: MLDS

Maloney and Yang (2003) used a direct method for estimating the maximum likelihood scale values,

$$L(\mathbf{\Psi}, \sigma) = \prod_{k=1}^{n} \Phi\left(\frac{\delta\left(\mathbf{q}^k\right)}{\sigma}\right)^{1-R_k} \left(1 - \Phi\left(\frac{\delta\left(\mathbf{q}^k\right)}{\sigma}\right)\right)^{R_k}$$

where

$$\Psi = (\psi_2, \psi_3, \ldots, \psi_{p-1})$$

$$\delta(\mathbf{q}^k) = |\psi_d - \psi_c| - |\psi_b - \psi_a|$$

Φ is the cumulative standard Gaussian (a probit analysis)

R_k is 0/1 if the judgment is lower/upper

$\psi_1 = 0, \psi_p = 1$ for identifiability,

leaving $p - 1$ parameters to estimate

Maloney LT, Yang JN (2003). “Maximum Likelihood Difference Scaling.” *Journal of Vision*, **3**(8), 573–585. URL http://www.journalofvision.org/3/8/5.

Estimation of Scale Values: MLCM

Ho, Landy & Maloney (2008) used a direct method for estimating the maximum likelihood scale values,

$$L(\mathbf{\Psi}, \sigma) = \prod_{k=1}^{n} \Phi\left(\frac{\delta\left(\mathbf{q}^k\right)}{\sigma}\right)^{1-R_k} \left(1 - \Phi\left(\frac{\delta\left(\mathbf{q}^k\right)}{\sigma}\right)\right)^{R_k}$$

where

$$\Psi = (\psi_2, \cdots, \psi_p, \chi_2, \cdots, \chi_q)$$

$$\delta(q^k) = (\psi^{b_1} + \chi^{g_1}) - (\psi^{b_2} + \chi^{g_2})$$

Φ is the cumulative standard Gaussian (a probit analysis)

R_k is 0/1 if the judgment is left/right image

$\psi_1 = \chi_1 = 0$ and $\sigma = 1$ for identifiability,

leaving $p + q - 2$ parameters to estimate

Estimation of Scale Values: MLDS

This problem can, also, be conceptualized as a GLM.

Each level of the stimulus is treated as a covariate in the design matrix, taking on values of 0 or ± 1,

depending on the presence of the stimulus in a trial and its weight in the decision variable.

	resp	S1	S2	S3	S4
1	0	4	8	2	3
2	1	2	3	6	11
3	1	2	6	7	10
4	0	4	11	1	2
5	0	9	11	7	8
6	0	7	10	1	3
		. . .			

p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}
0	1	−1	−1	0	0	0	1	0	0	0
0	1	−1	0	0	−1	0	0	0	0	1
0	1	0	0	0	−1	−1	0	0	1	0
1	−1	0	−1	0	0	0	0	0	0	1
0	0	0	0	0	0	1	−1	−1	0	1
1	0	−1	0	0	0	−1	0	0	1	0
					. . .					

For model identifiability, we drop the first column (fixing $\psi_1 = 0$ and $\sigma = 1$), *equal variance, Gaussian, signal detection model*.

The estimated scale is unique up to linear transformations.

Estimation of Scale Values: MLCM

The problem can also be conceptualized as a GLM.

Each level of the stimulus is treated as a covariate in the model matrix, taking on values of 0 or ± 1 in the design matrix,

depending on the presence of the stimulus in a trial and

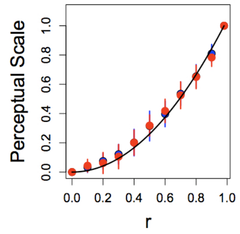
its weight in the decision variable.

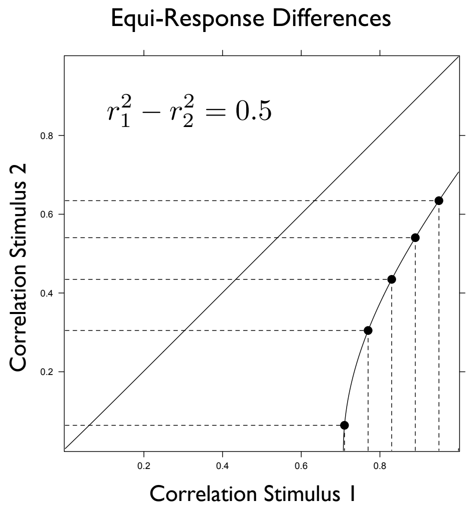
	Resp	G1	G2	B1	B2
1	1	3	4	4	3
2	1	3	5	4	2
3	0	1	1	1	4
4	0	2	3	1	2
5	0	1	4	3	4
6	1	1	5	5	2

p_1	p_2	p_3	p_4	p_5	q_1	q_2	q_3	q_4	q_5
0	0	1	−1	0	0	0	−1	1	0
0	0	1	0	−1	0	−1	0	1	0
0	0	0	0	0	1	0	0	−1	0
0	1	−1	0	0	1	−1	0	0	0
1	0	0	−1	0	0	0	1	−1	0
1	0	0	0	−1	0	−1	0	0	1

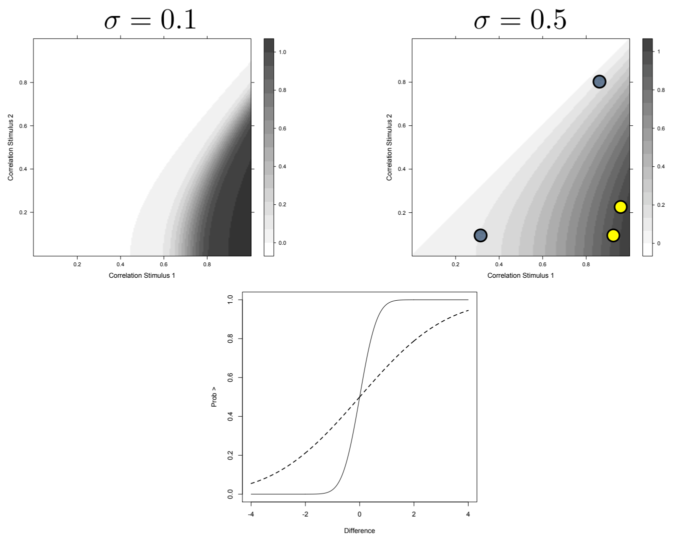
For model identifiability, we drop the first two columns along each dimension, fixing $\psi_1 = \chi_1 = 0$ and $\sigma = 1$.

$\psi(r) \approx r^2$

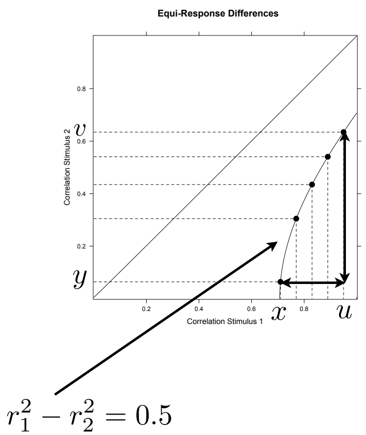




$r_1^2 - r_2^2 = 0.5$



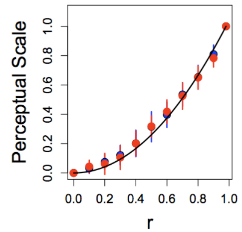
Indifference Curve: MLDS



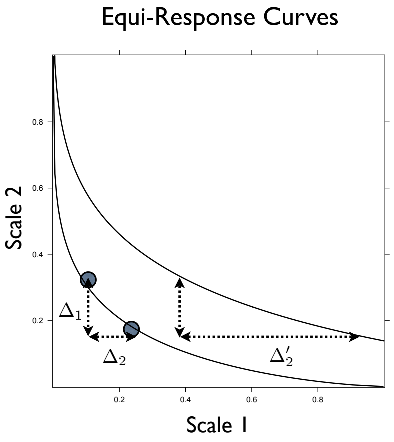
$(\psi_x - \psi_y) = (\psi_u - \psi_v)$

but also

$(\psi_x - \psi_u) = (\psi_y - \psi_v)$



Indifference Curves: MLCM



$$B_1 = B_2$$

$$\psi^b(b_1) + \chi^g(g_1) = \psi^b(b_2) + \chi^g(g_2)$$

$$\psi^b(b_1) - \psi^b(b_2) = \chi^g(g_2) - \chi^g(g_1)$$

$$\Delta_1 = \Delta_2 \approx \Delta_2'$$

The MLDS package

The MLDS package provides a modeling function, mlds(), that is essentially a wrapper for either glm() or optim(), and will enable estimation of the perceptual scale values, given a data frame with the previously described structure.

```
mlds(data, stimulus, method = "glm", lnk = "probit",
                opt.meth = "BFGS", opt.init = NULL,
                control = glm.control(maxit = 50000, epsilon = 1e-14),
                ... )
```

It outputs an S3 object of class ‘mlds’ which can be examined further using several method functions:

summary, plot, predict, fitted, logLik, boot, coef, vcov

The MLDS package

The data sets have class ‘mlds.df’ that inherits from ‘data.frame’. It differs in including two attributes, “stimulus” and “invord”.

```
> str(kk1)
Classes 'mlds.df' and 'data.frame':    330 obs. of  5 variables:
 $ resp: int  1 0 0 0 1 1 1 1 0 1 ...
 $ S1  : int  2 6 7 6 6 6 1 3 2 3 ...
 $ S2  : int  4 9 9 7 7 9 2 5 5 4 ...
 $ S3  : int  6 1 2 2 1 1 8 10 7 5 ...
 $ S4  : int  8 4 3 5 3 5 9 11 8 10 ...
 - attr(*, "invord")= logi  FALSE  TRUE  TRUE  TRUE  TRUE  
TRUE ...
 - attr(*, "stimulus")= num  0.0 0.1 0.2 0.3 0.4 ...
```

<u>stimulus</u> is a numeric vector of the physical stimulus levels

<u>invord</u> is a logical vector indicating whether on each trial the higher scale values were on the botton or top.

The MLCM package

The MLCM package provides a modeling function, mlcm(), that is essentially a wrapper for glm() and will enable estimation of the perceptual scale values, given a data frame with the appropriate structure.

```
mlcm(x,
    model = "add",
    whichdim = NULL,
    lnk = "probit",
    control = glm.control(maxit = 50000, epsilon = 1e-14),
    ...)
```

Default model is “additive”, but 2 others may be specified: “independent” (must specify `whichdim`) and “full”.

It outputs an S3 object of class ‘mlcm’ which can be examined further using several method functions:

summary, anova, plot, logLik, coef, vcov

The MLCM package

The data sets have class 'mlcm.df' that inherits from 'data.frame'.

```
> str(BumpyGlossy)
Classes 'mlcm.df' and 'data.frame':   975 obs. of  5 variables:
 $ Resp: Factor w/ 2 levels "0","1": 2 2 2 1 1 2 1 2 2 2 ...
 $ G1  : num  3 2 3 1 2 1 1 1 2 2 ...
 $ G2  : num  4 2 5 1 3 1 4 5 2 3 ...
 $ B1  : num  4 3 4 1 1 3 3 5 3 3 ...
 $ B2  : num  3 3 2 4 2 3 4 2 3 2 ...
```

Additive Model

```
> ( bg.add <- mlcm(BumpyGlossy) )

Maximum Likelihood Conjoint Measurement

Model:      Additive
Perceptual Scale:
       G      B
Lev1 0.000  0.000
Lev2 0.132  1.693
Lev3 0.185  2.947
Lev4 0.504  4.281
Lev5 0.630  5.275
```

Independent Model

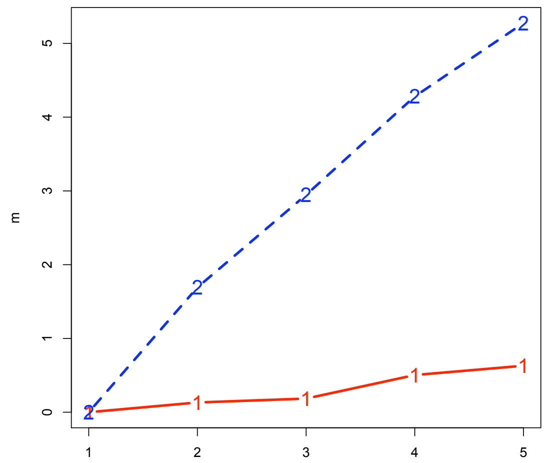
```
> ( bg.ind <- mlcm(BumpyGlossy, model = "ind", whichdim = 2) )

Maximum Likelihood Conjoint Measurement

Model:      Independence
Perceptual Scale:
    [,1]
B1  0.00
B2  1.66
B3  2.88
B4  4.16
B5  5.11
```

```
> anova(bg.ind, bg.add, test = "Chisq")
Analysis of Deviance Table

Model 1: resp ~ B2 + B3 + B4 + B5 - 1
Model 2: resp ~ (G2 + G3 + G4 + G5 + B2 + B3 + B4 + B5) - 1
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       971     500.12
2       967     476.48  4   23.635 9.452e-05 ***
```

We can also test a "full" model with 24 parameters!

```
> bg.full <- mlcm(BumpyGlossy, model = "full")

Model:    Full
Perceptual Scale:
   B1   B2   B3   B4   B5
G1 0.00 2.93 4.30 5.58 6.33
G2 1.20 2.95 4.07 5.55 6.73
G3 1.12 3.12 4.45 5.50 6.52
G4 1.96 3.45 4.33 5.93 6.88
G5 1.73 3.21 4.82 6.08 7.25

> anova(bg.add, bg.full, test = "Chisq")
Analysis of Deviance Table

Model 1: resp ~ (G2 + G3 + G4 + G5 + B2 + B3 + B4 + B5) - 1
Model 2: resp ~ (`G2:B1` + `G3:B1` + `G4:B1` + `G5:B1` + `G1:B2` + `G2:B2` +
    `G3:B2` + `G4:B2` + `G5:B2` + `G1:B3` + `G2:B3` + `G3:B3` +
    `G4:B3` + `G5:B3` + `G1:B4` + `G2:B4` + `G3:B4` + `G4:B4` +
    `G5:B4` + `G1:B5` + `G2:B5` + `G3:B5` + `G4:B5` + `G5:B5`) -
    1
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       967     476.48
2       951     461.53 16   14.947    0.5285
```

```
plot(bg.add, type = "b", col = c("red", "blue"), lwd = 3, cex = 1.5)
```

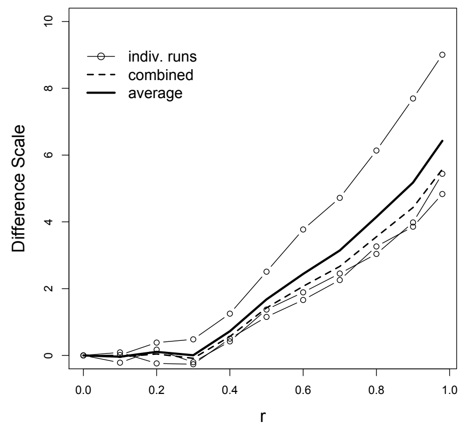


Mixed-effects models with MLDS (MLCM)

Three Strategies

1. Re-parameterize in terms of parametric decision variable
2. Normalize to common scale
3. Regression on estimated coefficients

Mixed-effects models with MLDS (MLCM): Re-parameterize in terms of decision variable



$$\Delta = \psi_d - \psi_c - \psi_b + \psi_a$$

re-parameterized as empirical decision variable:

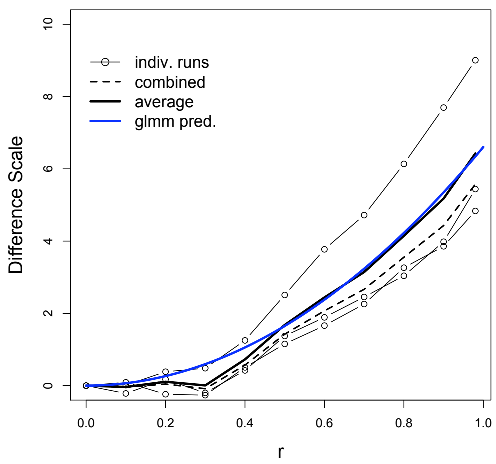
$$\mathcal{DV} = \rho_d^2 - \rho_c^2 - \rho_b^2 + \rho_a^2$$

then, fit GLMM

$$\Phi^{-1}(\mathsf{E}[Y]) = (\beta + b_i)\mathcal{DV},$$
$$b \sim \mathcal{N}(0, \sigma^2)$$

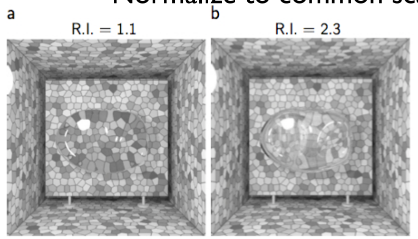
```
  resp S1 S2 S3 S4 Run    dv
1    1  2  4  6  8 kk1  0.16
2    0  6  9  1  4 kk1 -0.30
3    0  7  9  2  3 kk1 -0.25
4    0  6  7  2  5 kk1  0.04
5    1  6  7  1  3 kk1 -0.07
6    1  6  9  1  5 kk1 -0.23
             ...
kk.glmm <- glmer(resp ~ dv + (dv + 0 | Run) - 1,
   data = kk123, family = binomial("probit"))

summary(kk.glmm)

Generalized linear mixed model fit by the Laplace approximation
Formula: resp ~ dv + (dv + 0 | Run) - 1
   Data: kk123
   AIC   BIC logLik deviance
 650.8 660.6 -323.4    646.8
Random effects:
 Groups Name Variance Std.Dev.
 Run    dv   2.2752   1.5084
Number of obs: 990, groups: Run, 3

Fixed effects:
   Estimate Std. Error z value Pr(>|z|)
dv    6.604      0.962   6.865 6.65e-12 ***
```

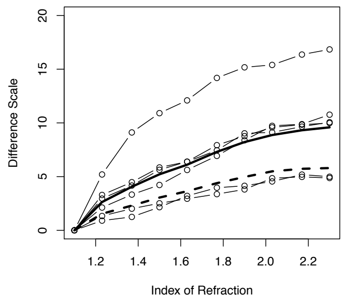
Mixed-effects models with MLDS (MLCM): Re-parameterize in terms of decision variable




```
> coef(kk.glmm)
$Run
          dv
kk1 5.433941
kk2 8.459360
kk3 5.761285
```

Mixed-effects models with MLDS (MLCM):
Normalize to common scale





Experiment of Fleming, Jäkel and Maloney.
Perception of transparency as a function of rendered index of refraction

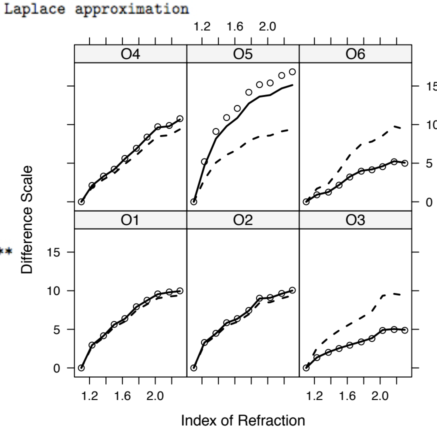
No simple functional description of relation because of kink in curve

Use each individual's scale value to compute decision variables and fit GLMM to these value; normalizes out individual shape differences.

$$\mathrm{DV}_o = \hat{\psi}_{d,o} - \hat{\psi}_{c,o} - \hat{\psi}_{b,o} + \hat{\psi}_{a,o}$$

Mixed-effects models with MLDS (MLCM):
Normalize to common scale

```
Generalized linear mixed model fit by the Laplace approximation
Formula: resp ~ DV + (DV + 0 | Obs) - 1
   Data: Transparency
  AIC  BIC logLik deviance
 1485 1497   -741     1481
Random effects:
 Groups Name Variance Std.Dev.
 Obs    DV   13.7     3.69
Number of obs: 2520, groups: Obs, 6

Fixed effects:
   Estimate Std. Error z value Pr(>|z|)
DV     9.39       1.57       6    2e-09 ***
```



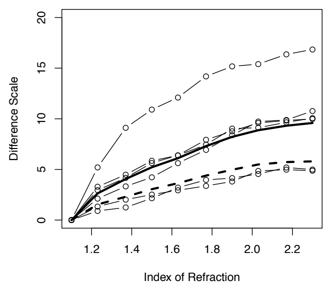
Mixed-effects models with MLDS (MLCM):
Regression on estimated coefficients

For this approach we use lmer and fit the coefficients as a function of the stimulus level using MLDS directly.

$$\hat{\psi}(S) \sim (\beta_1 + b_1)S + (\beta_2 + b_2)S^2 + \cdots + \epsilon$$

By taking the log of the coefficients, we transform the multiplicative effect to additive.
We use polynomials to fit the fixed effect but also to model random differences in the shapes of the function across observers

$$\log(\hat{\psi}(S)) \sim (\beta_0 + b_0) + (\beta_1 + b_1)S + (\beta_2 + b_2)S^2 + \cdots + \epsilon$$



Mixed-effects models with MLDS (MLCM):
Regression on estimated coefficients

$$\log(\hat{\psi}(S)) \sim (\beta_0 + b_0) + (\beta_1 + b_1)S + (\beta_2 + b_2)S^2 + \cdots + \epsilon$$

First, test random effects:

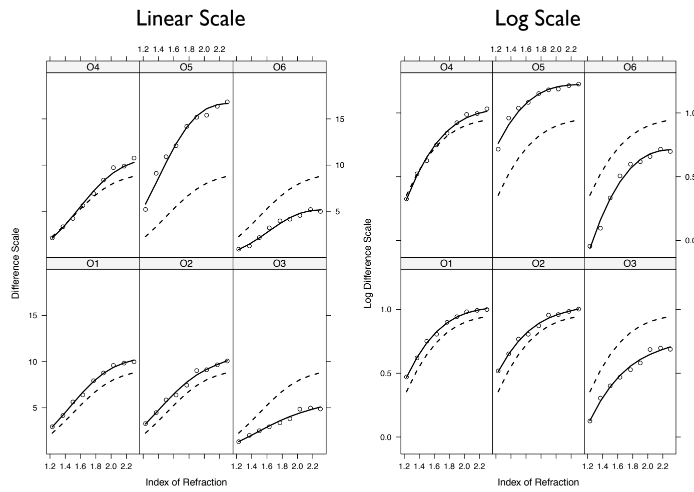
```
> P3 <- lmer(logDS ~ poly(Stim, degree = 6) +
+         (Stim + I(Stim^2) + I(Stim^3) | Obs), Tr.df)
> P2 <- lmer(logDS ~ poly(Stim, degree = 6) +
+         (Stim + I(Stim^2) | Obs), Tr.df)
> P1 <- lmer(logDS ~ poly(Stim, degree = 6) +
+         (Stim | Obs), Tr.df)
> P0 <- lmer(logDS ~ poly(Stim, degree = 6) +
+         (1 | Obs), Tr.df)
> anova(P0, P1, P2, P3)

Data: Tr.df
Models:
P0: logDS ~ poly(Stim, degree = 6) + (1 | Obs)
P1: logDS ~ poly(Stim, degree = 6) + (Stim | Obs)
P2: logDS ~ poly(Stim, degree = 6) + (Stim + I(Stim^2) |
P2:     Obs)
P3: logDS ~ poly(Stim, degree = 6) + (Stim + I(Stim^2) +
P3:     I(Stim^3) | Obs)
   Df  AIC  BIC logLik Chisq Chi Df Pr(>Chisq)
P0  9 -126 -108   71.9
P1 11 -157 -135   89.7 35.45      2      2e-08 ***
P2 14 -162 -135   95.2 11.19      3      0.011 *
P3 18 -158 -123   97.2  3.94      4      0.414
```

Mixed-effects models with MLDS (MLCM):
Regression on estimated coefficients

Then, test fixed effects:

```
> P2.2 <- lmer(logDS ~ poly(Stim, degree = 2) +
+          (Stim + I(Stim^2) | Obs), Tr.df)
> P2.3 <- lmer(logDS ~ poly(Stim, degree = 3) +
+          (Stim + I(Stim^2) | Obs), Tr.df)
> P2.4 <- lmer(logDS ~ poly(Stim, degree = 4) +
+          (Stim + I(Stim^2) | Obs), Tr.df)
> P2.5 <- lmer(logDS ~ poly(Stim, degree = 5) +
+          (Stim + I(Stim^2) | Obs), Tr.df)
> anova(P2, P2.5, P2.4, P2.3, P2.2)

Data: Tr.df
Models:
P2.2: logDS ~ poly(Stim, degree = 2) + (Stim + I(Stim^2) |
P2.2:     Obs)
P2.3: logDS ~ poly(Stim, degree = 3) + (Stim + I(Stim^2) |
P2.3:     Obs)
P2.4: logDS ~ poly(Stim, degree = 4) + (Stim + I(Stim^2) |
P2.4:     Obs)
P2.5: logDS ~ poly(Stim, degree = 5) + (Stim + I(Stim^2) |
P2.5:     Obs)
P2: logDS ~ poly(Stim, degree = 6) + (Stim + I(Stim^2) |
P2:     Obs)
     Df  AIC  BIC logLik Chisq Chi Df Pr(>Chisq)
P2.2 10 -163 -143   91.5
P2.3 11 -167 -145   94.5  6.03      1      0.014 *
P2.4 12 -166 -142   95.1  1.24      1      0.266
P2.5 13 -164 -138   95.1  0.04      1      0.838
P2   14 -162 -135   95.2  0.22      1      0.637
```

Mixed-effects models with MLDS (MLCM):
Regression on estimated coefficients



- *Difference Scaling* and *Conjoint Measurement* are psychophysical techniques that permit estimation of interval perceptual scales by maximum likelihood
- The two approaches are implemented in R packages MLDS and MLCM, respectively, on CRAN.
- We can introduce mixed-effects into the MLDS and MLCM (not shown) models using the lme4 package (and perhaps others).

Thank you.