NAFTA Developmental Neurotoxicity Guidance Document

Startle ResponseTesting

Kevin M. Crofton, PhD
R3Fellows, LLC
Durham NC USA

orcid.org/0000-0003-1749-9971

EFSA DNT Lecture Series
December 2020

Outline

- Lecture Series on NAFTA Guidance Continues
  - Past lectures
    - NAFTA Background and FOB & Clinical Observations
    - Motor Activity
  - This time – Startle Response,
- Guideline Requirements
- What is the startle response?
  - Basic biology of the startle response
  - Common methods used
  - Important features that influence the startle response
  - Ontogeny of the startle response
  - Variability
- Interpretation
  - Examples of control data
  - Examples of test data – the good, the bad, and the ugly
  - Body weight as a confound??
- Summary

Goals of the NAFTA Document

- Improve understanding of DNT guideline studies
  - Appropriate conduct of behavioral tests
  - Biological significance of endpoints
- Increase consistency in interpretation and assessment of outcomes
- Focus on behavioral tests
  - Observations
  - Motor activity
  - **Auditory startle response**
  - Learning and memory
- Neuropathology/morphometrics not included
  - Several excellent peer-reviewed publications already available

Courtesy V. Moser

DNT Guidelines

Testing Requirements for Motor Activity

	EPA 870.6300	OECD 426	OECD 443
Test species	Rat	Rat	Rat
Exposure	GD6 to weaning	GD6 to weaning	2 weeks pre-mating to weaning
Motor activity	Preweaning ontogeny and adult	Preweaning and adult	Preweaning and adult
Neuromotor ontogeny	None	Preweaning	None
Functional/Clinical observations	Throughout	Throughout	Adult
Auditory startle response	**Weaning and adult**	**Weaning and adult**	**Weaning**
Learning and memory	Weaning and adult	Weaning and adult	None
Neuropathology and morphology	Weaning and adult	Weaning and adult	Adult

Courtesy V. Moser

Requirements (con't)

EPA 870.6300* – Auditory Startle Habituation

- Two ages – "around the time of weaning and around day 60"
- Does not state that same animals must be used
- 50 test trials (5 blocks of 10 trials on each day)
- Treatment groups should be:
  - Counter-balanced across test devices
  - Counter-balanced across test times
- Prepulse inhibition is not a requirement

USEPA, 1998

Requirements (con't)

OECD 426 – Motor and Sensory Function

- Two Ages – "should be examined in detail at least once for the adolescent period and once during the young adult period (*e.g.*, PND 60-70)."
- Does not specifically require startle
- Provides references to published papers on auditory startle habituation
- Most studies follow EPA guidelines

OECD 443 – Auditory Startle Test

- **One age only –**PND 24 (±1 day)
- Counterbalancing across day
- 50 test trials (5 blocks of 10 trials on each day) *with test conditions optimized to produce intra-session habituation.*
- Procedures should be consistent with OECD 426

OECD, 2007; 2012

Why measure startle in DNT studies

- A simple measurement of the integrity a sensory-evoked motor reflex
  - No training required
  - Rapid objective automated methods
  - Extrapolation potential – occurs in all mammals
- Repeated trials allows estimates of habituation
- Known developmental ontogeny
- Decades of research show it be capable of detecting many chemical stressors

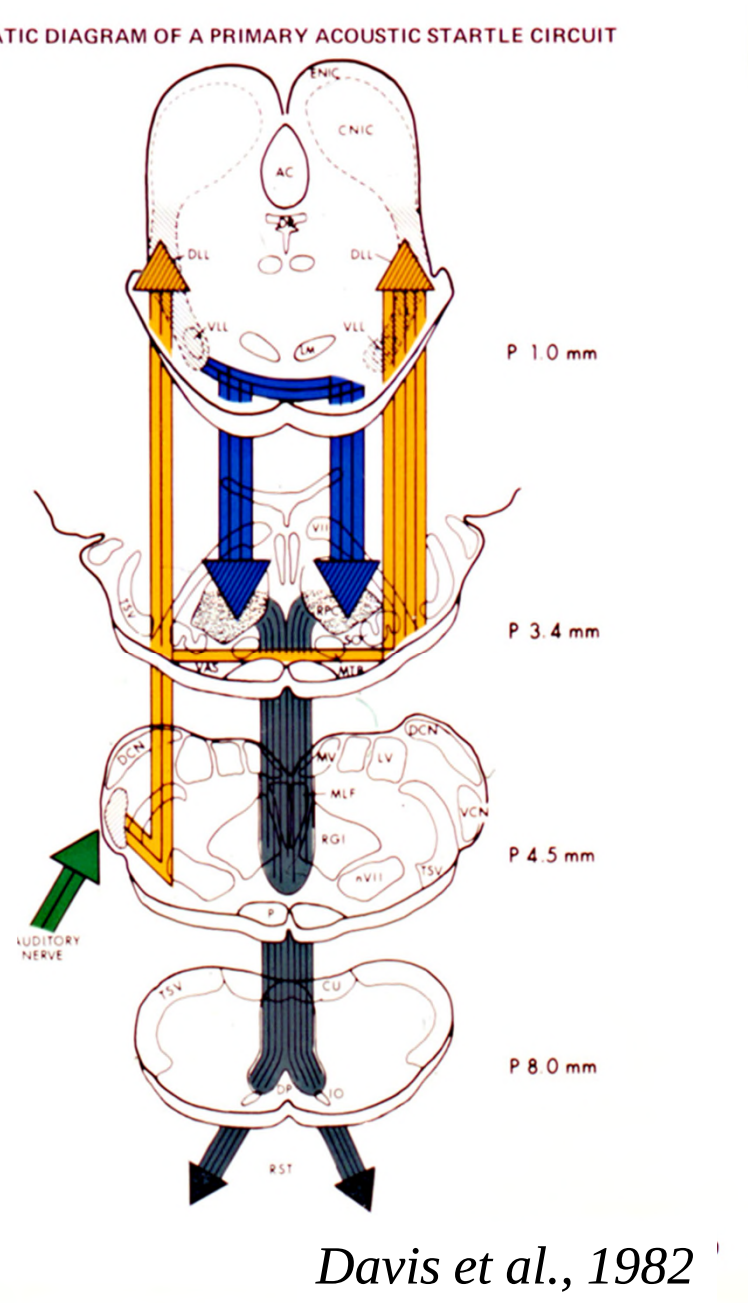
What is the startle response?

- "The startle response consists of a characteristic sequence responses elicited by a sudden intense stimulus" (Davis, 1984). It is characterized by, the rapid contraction and extension of skeletal muscles.
  - Best induced by **auditory stimuli**, but also visual and tactile stimuli
  - The amplitude is a graded response well suited for quantitative analyses

The light flash induced startle response of a nine-banded armadillo can reach up to one meter. By Bianca Lavies, 1982 National Geographic Society.

Basic Anatomy of the Auditory Startle Response

- Cochlea
- Auditory nerve
- Brainstem nuclei
- Spinal tracts
- Neuromuscular junction
- About 8-10 msec from stimulus onset to neuromuscular junction
- No higher cortical processing is required



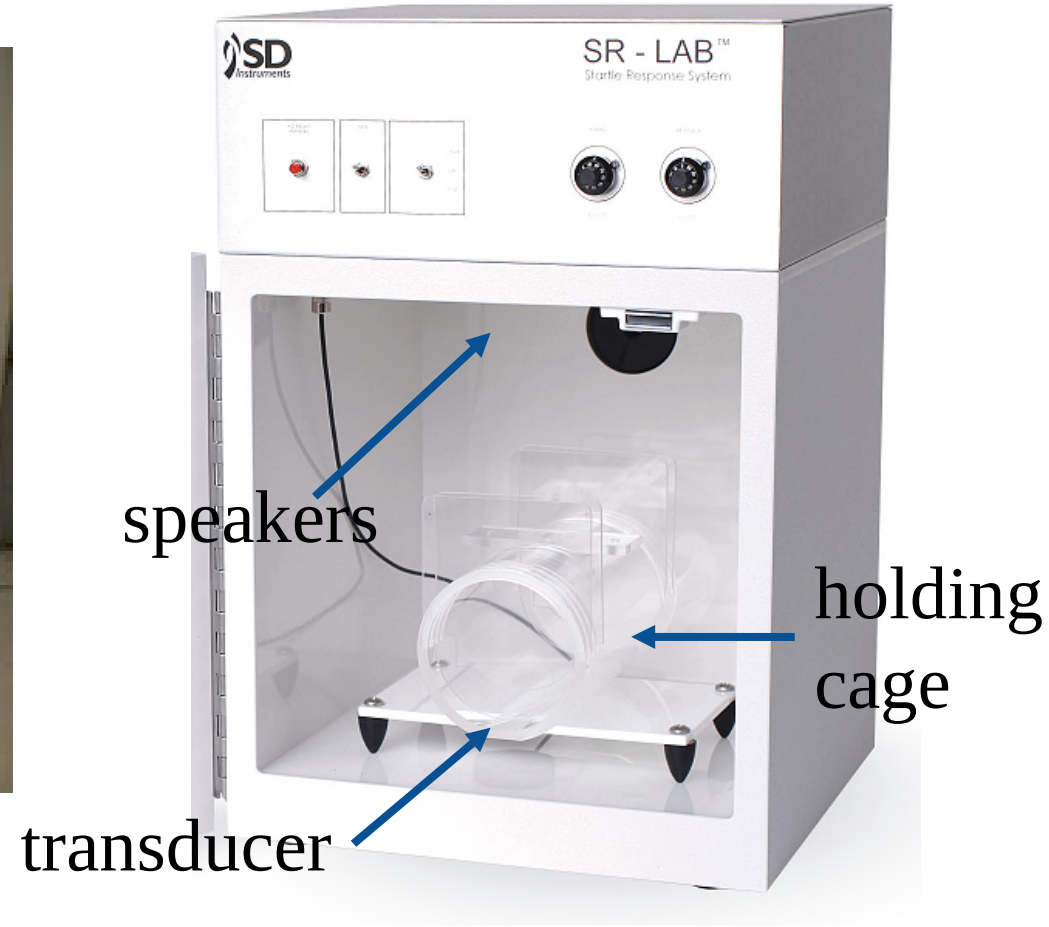
Davis et al., 1982

Note: Koch and colleagues have expanded on this pathway (see Koch, 1999)

Common Methods

Most commercial methods include:

- Sound attenuated test chambers, auditory stimulus systems, speakers, animal chambers (cages), response measurement platform
- Common to have multiple chambers for more rapid testing



Factors that may influence startle measurements

Organismal Factors

- Age, Sex, Species and Strain*, Previous Experience, Food Deprivation (same as all behavioral methods)

Most Important Experimental Factors**

- **Detection method**
- **Acclimation time to chamber**
- **Noise** – both in chamber and external
- Number of trials
- Circadian rhythms
- Counter-balance across test chambers

** Strain effects: Glowa and Hansen, 1994*

***Reviews: See Hoffman and Ison, 1980; Eaton, 1984.*

Calibration is Important

- Speakers must be calibrated for all auditory stimuli
  - Both background noise and startle stimulus
    - Background noise commonly about 40dB
    - Startle stimulus need to be about 115-120 dB to reliability elicite a startle response)
    - Must be calibrated for rat not human hearing*
- Detection platforms must be calibrated against a standard
- For systems with multiple test chambers all must be calibrated to the same standards!

**** See NAFTA 2016 p.C-7***

Common Detection Methods

- **Accelerometer**
  - Measures changes in force on the platform
  - Calibrated with vibration equipment
  - Normal output in voltage or response units
  - Harder to calibrate across equipment companies

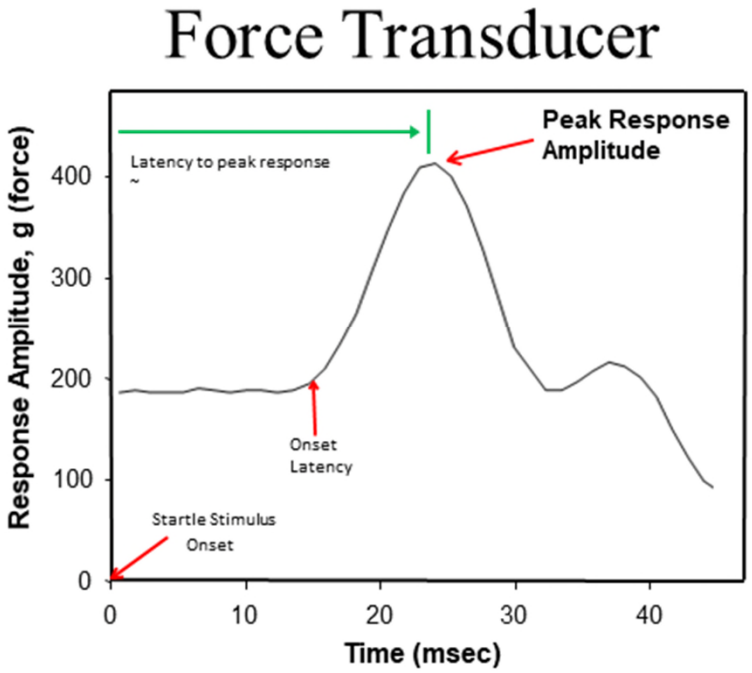
- **Force transducer**
  - Measures the force applied to the platform
  - Calibrated with standardized weights
  - Output in grams (force)
  - Standardize across platforms and companies
  - More expensive

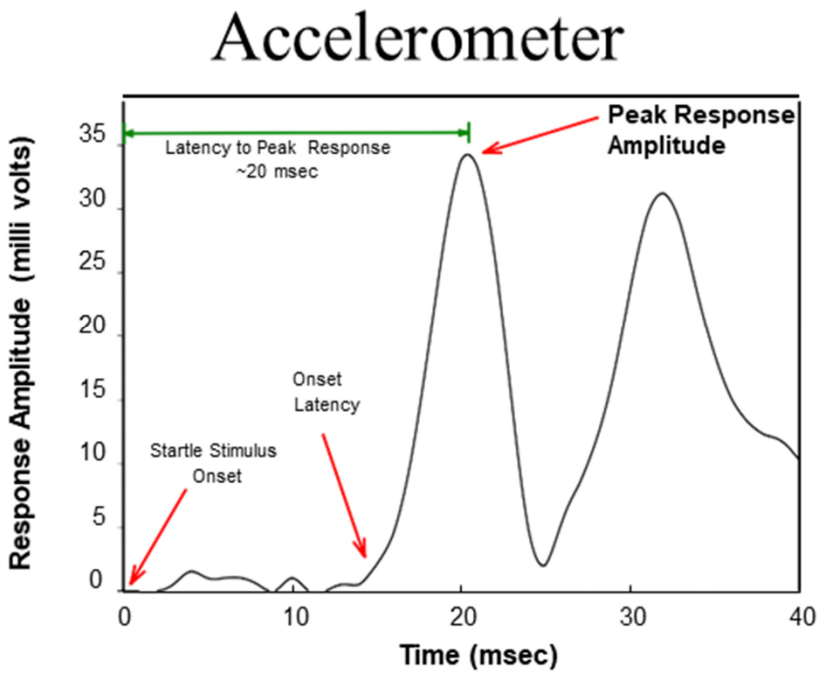
- **Video systems - new**
  - Not included here due to no use in 426 studies
  - For an example see Pantoni et al (2020)

Response Measurement (con't)

Response vs time for force transducer and accelerometer

- Both measure latency and amplitude
- Force transducer: body weight estimate, calibrated force in grams, latency to peak force
- Accelerometer: force (millivolts, arbitrary units, Newtons), no body weight, latency to peak acceleration not peak of force
- Both are good – just understand the differences





Modified from NAFTA (2016)

Response Amplitude and Latency

- Amplitude
  - Commonly reported as the peak
  - Sometimes reported as average across the entire response time

- Latency
  - Reported as both onset and latency to peak
  - Onset should be in the 10-14 msec range
  - Latency to peak is correlated with amplitude and device (usually longer time to first peak with force transducer)
  - Very rare to find effect on latency without a change in amplitude

Terminology Issues

Variable names are very variable ☺

Variable names found in 45 DNT studies

- 13 different names for amplitude and latency
- 3 were undefined in the report
- Peak amplitude and average amplitude are not the same

Make sure you know what is being reported

Table 2. Common reporting terms for startle response data and biological measurement interpretation

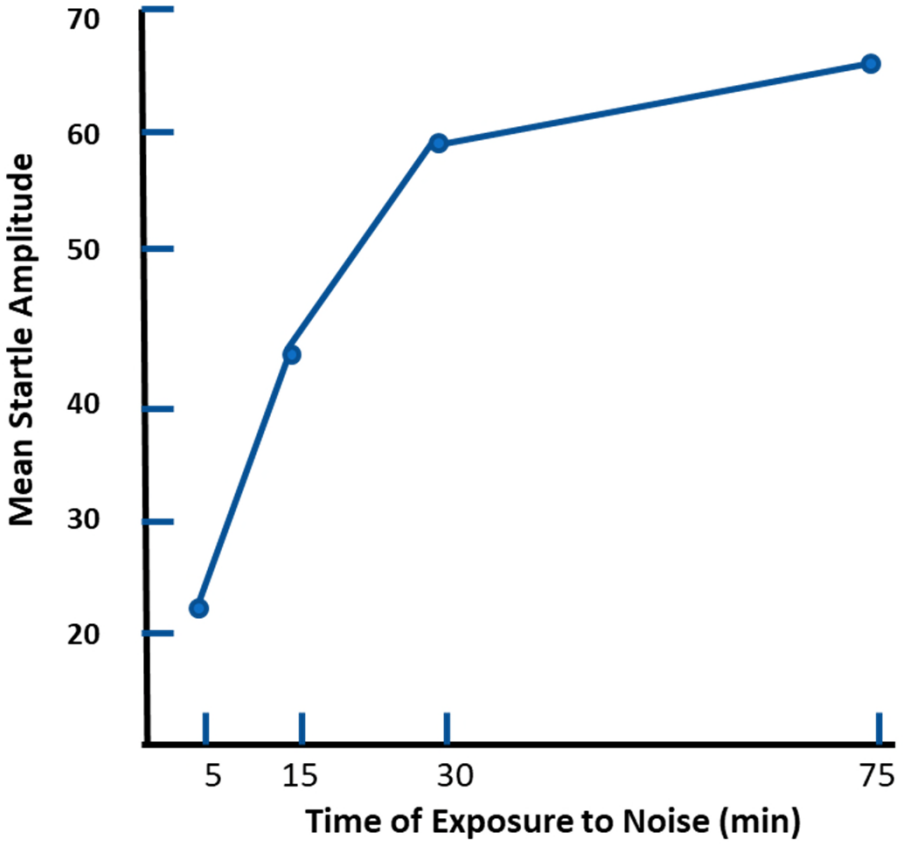
Reported Term	Measurement
amplitude (Newtons)	Amplitude
amplitude (volts)	Amplitude
maximum amplitude (no units)	Amplitude
maximum amplitude of the response (mV)	Amplitude
maximum impulse (arbitrary units)	Amplitude
maximum response (Vmax)	Amplitude
maximum response amplitude (Vmax)	Amplitude
maximum startle (voltage)	Amplitude
mean peak startle (no units)	Amplitude
peak amplitude (grams)	Amplitude
peak response amplitude (v)	Amplitude
reflex peak amplitude (V)	Amplitude
response (g)	Amplitude
latency (msec)	Latency
latency (Tmax)	Latency
latency to maximum response (msec)	Latency
latency to peak (msec)	Latency
latency to peak of the response (msec)	Latency
latency to peak response	Latency
mean time to maximum amplitude	Latency
peak latency (msec)	Latency
reflex latency (msec)	Latency
time of maximum startle (msec)	Latency
time to maximum response (Tmax)	Latency
time to maximum startle (msec)	Latency
time to peak amplitude (msec)	Latency
percent inhibition	Sensory inhibition
prepulse inhibition	Sensory inhibition
average response (grams)	Undefined in report
average response amplitude (mV)	Undefined in report
response duration (msec)	Undefined in report

NAFTA, 2016

Acclimation & Background Noise Interaction

- Most test paradigms use background noise exposure during a 5-10 min acclimation to the test chambers to decrease variability in initial trials
- This interacts with amplitude of response
- Needs to be consistent for all animals

Important variable to track when reviewing studies

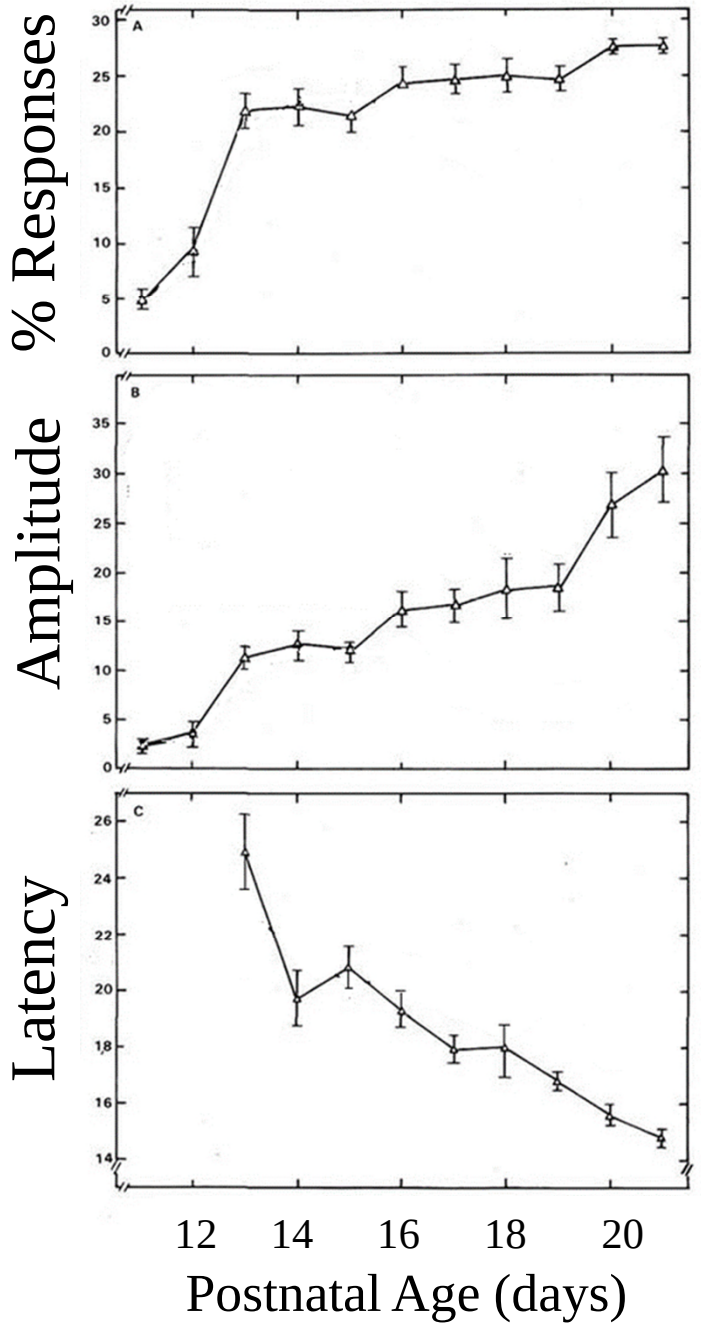


Time of exposure to background noise prior to starting testing

(redrawn from Davis 1974)

Ontogeny of the ASR in the rat

- Ontogeny is linked to maturation of the hearing and musculature during the second postnatal week
- Response incidence increases to 95% by PND21
- Latency decreases showing maturing response
- Amplitude increase with age and continues to mature beyond PND21



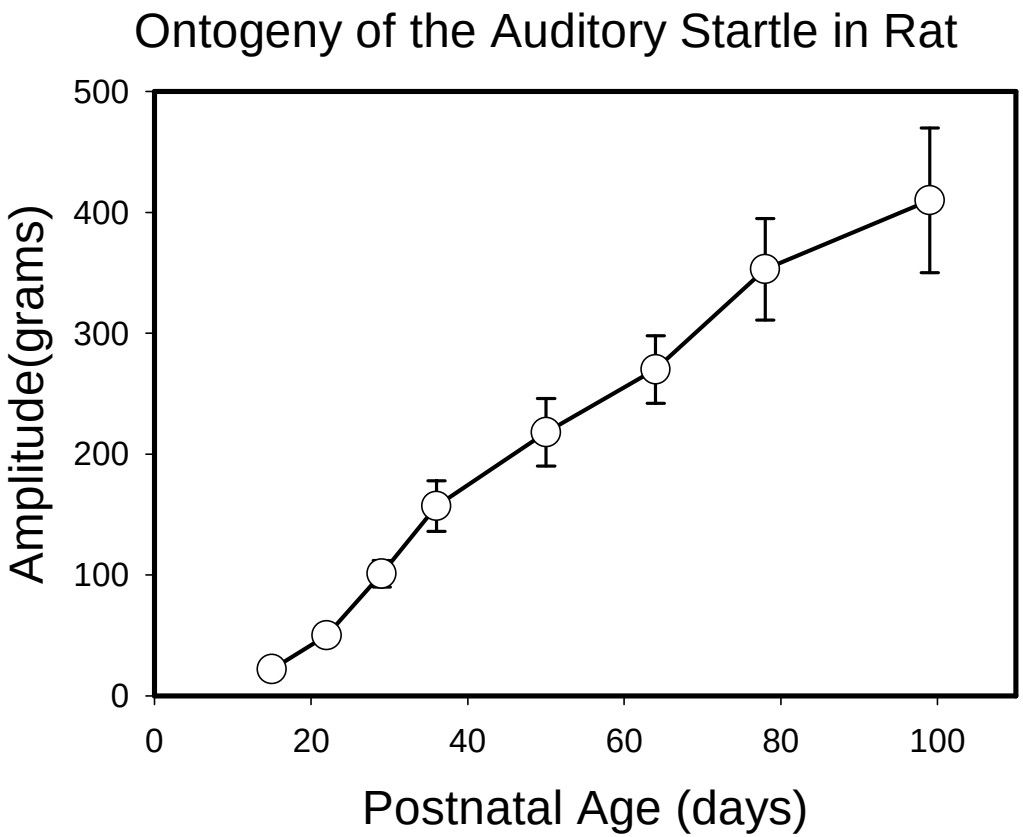
Modified from Sheets et al. 1988

The Development of the Auditory Startle Response (con't)

- Amplitude continues to rise as muscle mass increases through to adult ages

- The absolute value of the startle response should be higher in adult compared to weanling animals

Important variable to track when reviewing studies



Redrawn from Dean et al. 1990

Variability

Factors that influence data variability (same for most behavioral tests)

- **Intrinsic Variability**
  - "Biological" differences among individuals: Inherent to the parameter and test system
- **Extrinsic Variability**
  - Environment sources (experimental method & study design, etc)
  - These are variables that the investigator is able to control at all stages of the study

* *CV = standard deviation / mean*

** *CV was lowest in published papers from research labs*

Variability

- **Raffaele et al (2008)** is a great review on variability from DNT guideline study with a case-study focused on startle data
  - Reviewed startle data from both literature and fine commercial DNT tesing labs
  - Coefficient of Variation (CV)* varied from 20-110%*
    - *CV was lowest in published papers from research labs*
  - Higher variability at PND 22-24 compared to adults
  - High variability is usually a result a lack of control of experimental variables

** CV = standard deviation / mean*

Data Checks and Interpretations

- **Control group**
  - Amplitudes should be age-appropriate
    - Adult animals should be higher than weanling animals
  - Habituation should be present to some extent at PND24?
  - Variability should not be excessive and decline with age (see Raffaele et al. 2008)
  - These indicate that the testing lab has good experimental control of the method

Data Checks and Interpretations (con't)

Historical control and positive control data

- Historical control data
  - demonstrate reliability of techniques
  - Allows comparison to identify possible "outliers"

- Positive control data should be available & informs interpretation
  - Show lab competency to detect chemical induced changes
  - Show sufficiency of technical personnel
  - Defines the dynamic range
    - e.g., does a large dose of a sedative like compound or hearing loss cause a 25% decrease or a 90% decrease

Data Checks and Interpretations (con't)

- **Statistics (briefly)**
  - Within session testing (habituation) is a repeated measure!
  - Multiple testing of the same animal is a repeated measure!
  - Litter as a statistical unit.
  - SEX – must be included in the stats model
    - If not, there should be no conclusions of sex-dependent effects

Data Checks and Interpretations (con't)

- **Statistical vs biological significance**
  - There needs to be balance in determination of an effect
  - From NAFTA Guidance:

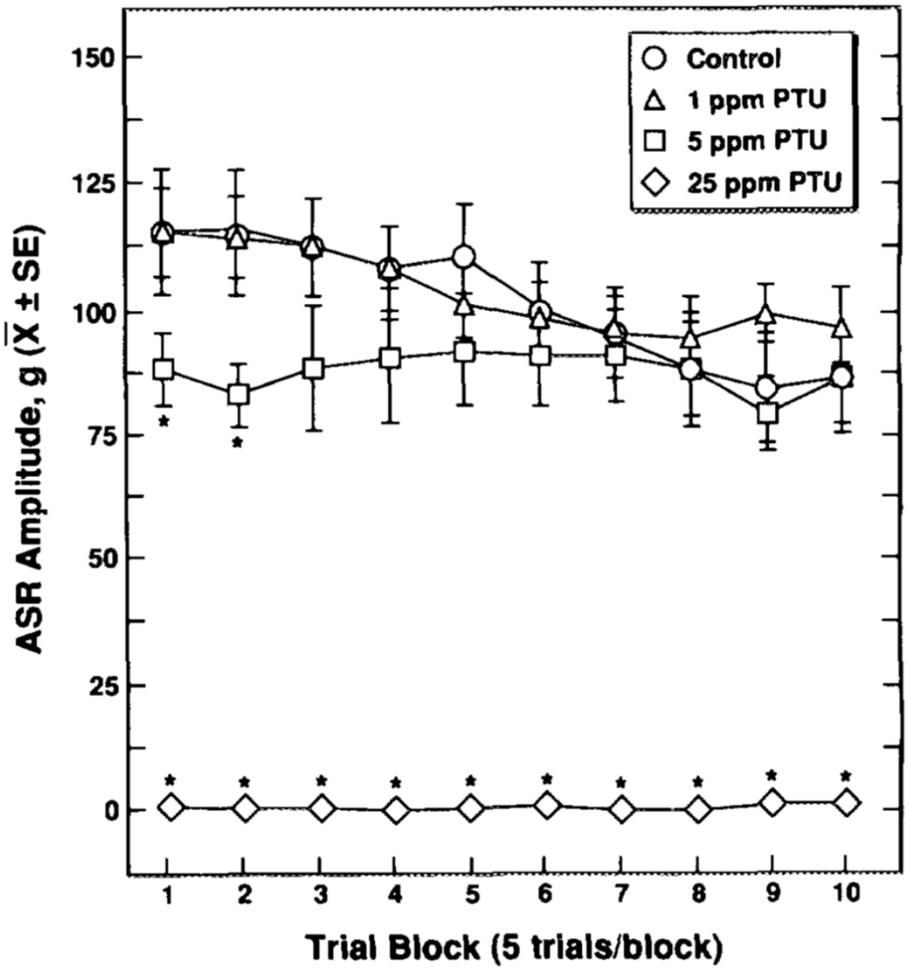
    "guideline DNT studies are conducted to screen chemicals for possible adverse effects on the developing nervous system and are often the only study examining all of these endpoints. *Thus, consideration of a higher false positive rate rather than a lower false negative rate may be a more conservative approach in some case*s.
- Large or excessive variance can occur due to poor control over testing and increases possible false negatives
  - Compare to within lab results (historical control) or other labs using similar equipment

Example Positive Control - PTU

- Propylthiouracil (PTU) in water supply from GD18 – PND21

- PND24 - Startle response
  - No habituation in mid dose
  - Abolished response in high dose animals that have severe hearing loss.

- Demonstrates dynamic range for decreases and impact on habituation



Goldey et al 1995

See also Crofton et al., 2004; Marable and Maurissen, 2004

Common Problems in Data Interpretation*

Table 6. Potential interpretation of some of the most common problems for startle results found in DNT study reports.

Data Problem		**Potential interpretations**
1	Lack of age dependent increase in control responses (e.g., PND21 vs PND60)	If not explained by equipment issues (e.g., sensitivity gain changes between age groups), then treatment-related effects, or lack thereof, may not be interpretable
2	Lack of sex dependent differences in control responses of adult rats	If not explained by equipment issues (e.g., sensitivity gain changes), then treatment-related effects, or lack thereof, may not be interpretable
3	Lack of evidence of habituation at young age (around weaning)	Animals may have been tested at too young an age, check historical and positive control data for habituation at that specific age. If age is not the cause, see #4 below
4	Lack of evidence of habituation at adult age	Lack of habituation suggests improper experimental conduct or data analysis. Verify that data were analyzed so that habituation was tested directly. If data were analyzed appropriately then the data should not be used for decisions based on habituation. Main effects of treatment may still be useful.
5	Significant effects on latency with no impact on amplitude	Since effects on latency without changes in amplitude are extremely rare and may not be biologically reasonable, efforts should be taken to ensure that testing equipment and data algorithms are correctly identifying response peak times.
6	Significant effects on startle amplitude in only one sex	Lack of effects in both sexes is not a valid reason for dismissing treatment-related effects. Chemicals that interfere with some endocrine systems during development may have sex-selective effects.
7	Significant effects of treatment only at the low or middle dose in one sex.	If sex is not included in the model, then reanalyze data including sex. If data re-analyses reveal no sex-by-treatment interaction and no main effect of treatment, then the reported effect was likely a false positive.
8	Dose-related effects of increasing magnitude with a lack of statistical significance coupled with excessive variation	Large effects on the amplitude of the startle response may be masked by excessive variability. Expert judgment is required to determine the biological significance in light of the lack of statistical significance. Check historical and positive control data. Check to ensure large variability is not due to extreme values for few test animals.

**Please see NAFTA, 2016 for details*

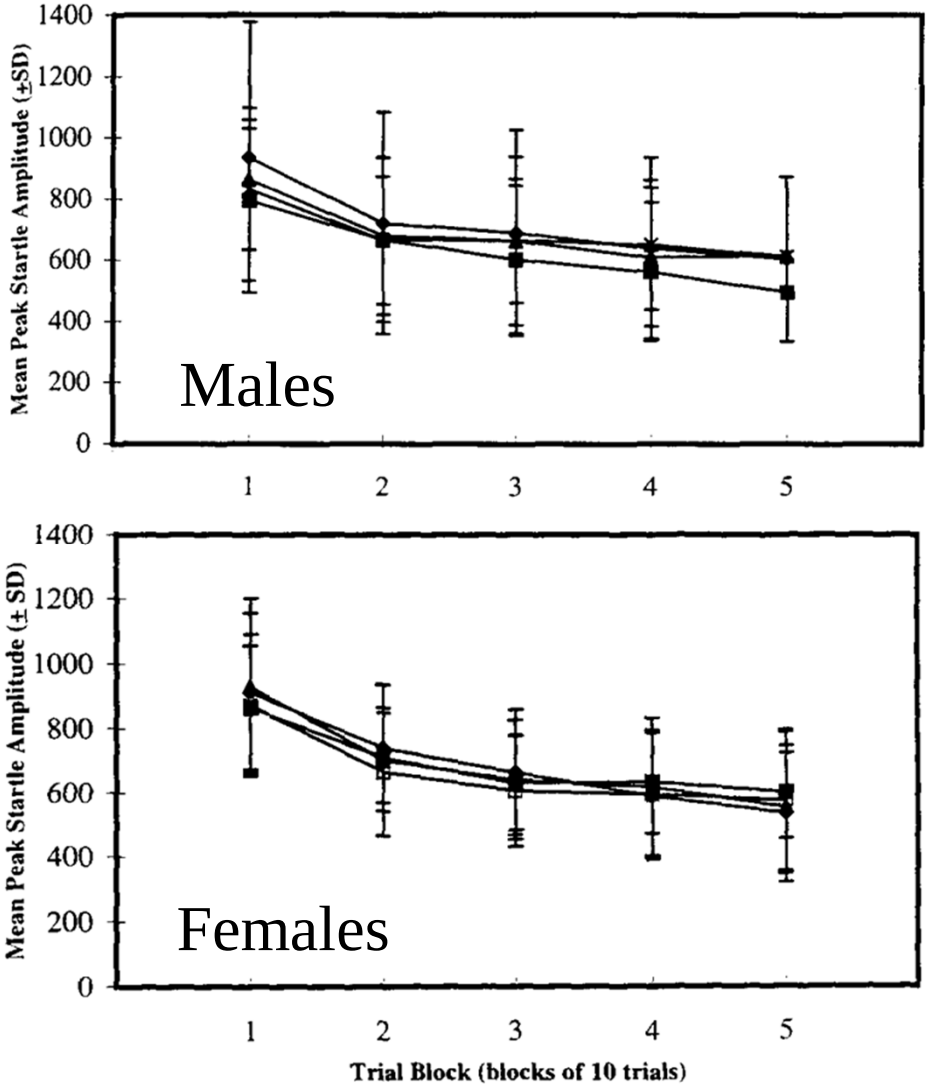
Examples – Good Report

- Company X submits DNT with the following
  1. Extensive methods section description of testing protocol
  2. Positive control data showing ability to detect increases and decreases in ASR (in adults)
  3. Historical controls from other DNT guideline studies

Examples (con't)

Pesticide X - PND22

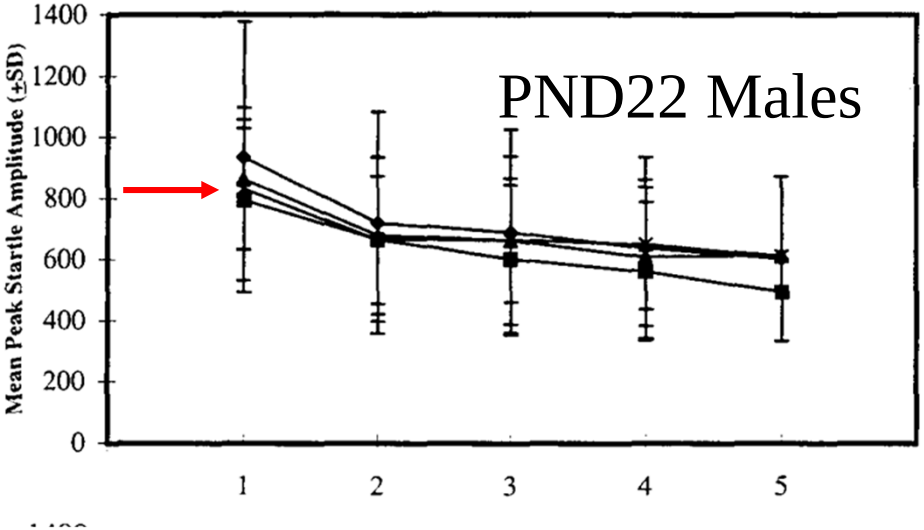
- Shows some habituation
- Variability is okay. CV=~40% in controls
- standard deviation on plots).
- Clearly no treatment effects
- No difference in amplitude between males and females (okay at this early age)*

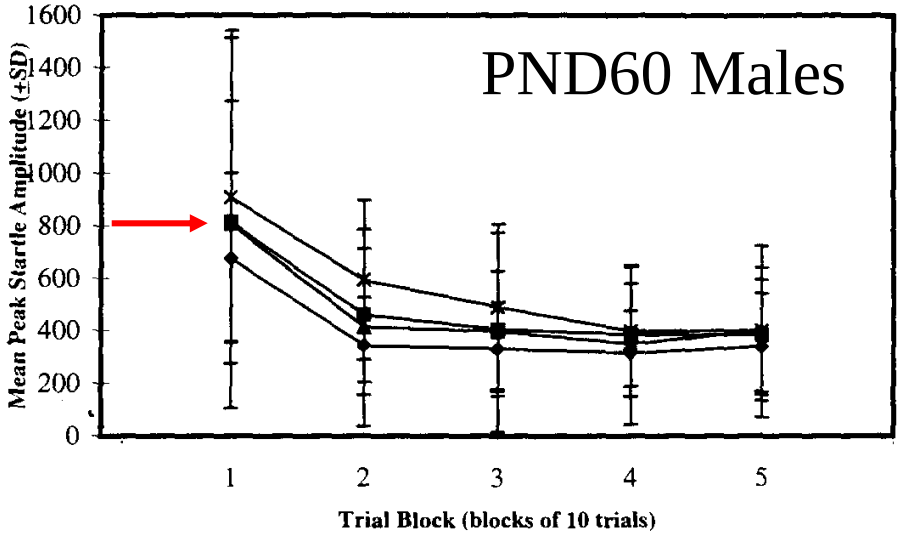


Examples – Problem??

Pesticide X – PND22 vs 64

- Still no treatment effects
- Control amplitude in adult is the same as on PND22? (red arrows)
- Careful reading and background on method used show this was due to change in calibration





Example – Problems

Guideline Study - Compound S Study

Test Age	Control	Low	Med	High
Females				
PND 20				
Vmax (mv)	181.2±50.8	182.5±54.1	160.6±47.9	129.0±56.6
Tmax (ms)	22.7±2.6	24.0±2.6	24.6±4.6	25.3±2.0
Vave (mv)	39.5±10.0	37.9±10.5	34.4±10.9	27.4±12.5
PND 60				
Vmax (mv)	78.2±36.1	78.7±30.7	79.6±40.4	80.1±37.0
Tmax (ms)	34.3±4.4	32.9±4.7	33.9±4.5	30.9±4.7
Vave (mv)	16.7±7.6	17.3±7.8	16.2±7.0	16.5±6.7

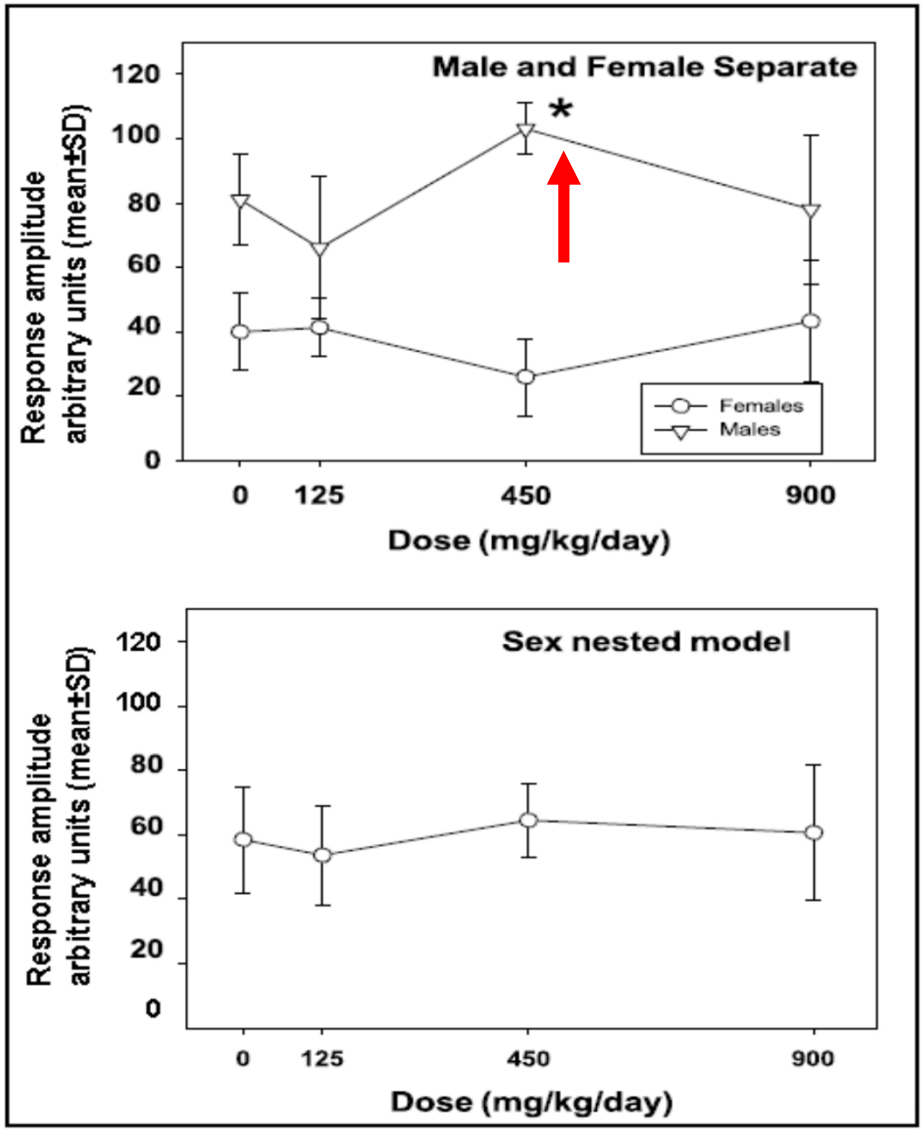
- Why is amplitude for PND 60 lower than PND20?
- Why is latency (Tmax) higher for the smaller amplitude?
- Methods section: accelerometer was used. But fails to document how calibrations were performed
- Historical controls were not provided.

Request info on calibration procedure and historical and positive control data.

Examples – Sex Dependent Effect?

Startle data from adult testing

- Report states statistically significant effect in males only at middle dose. Sexes were analyzed separately
- When data was re-analyzed with sex in model there was no main effect of treatment nor a any significant interaction of treatment and sex



Body Weight and ASR

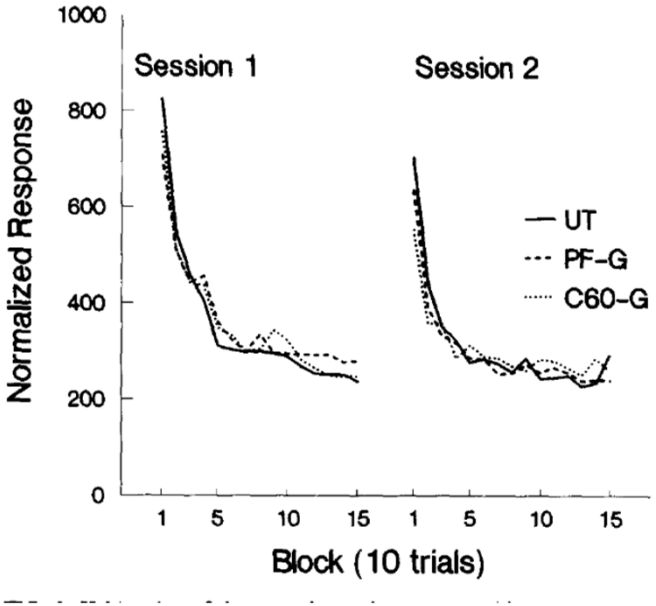
Fact: Body weight impacts startle response as animals grow and become larger

Fiction: Body weight should always be used as a covariate when treatment alters body weight

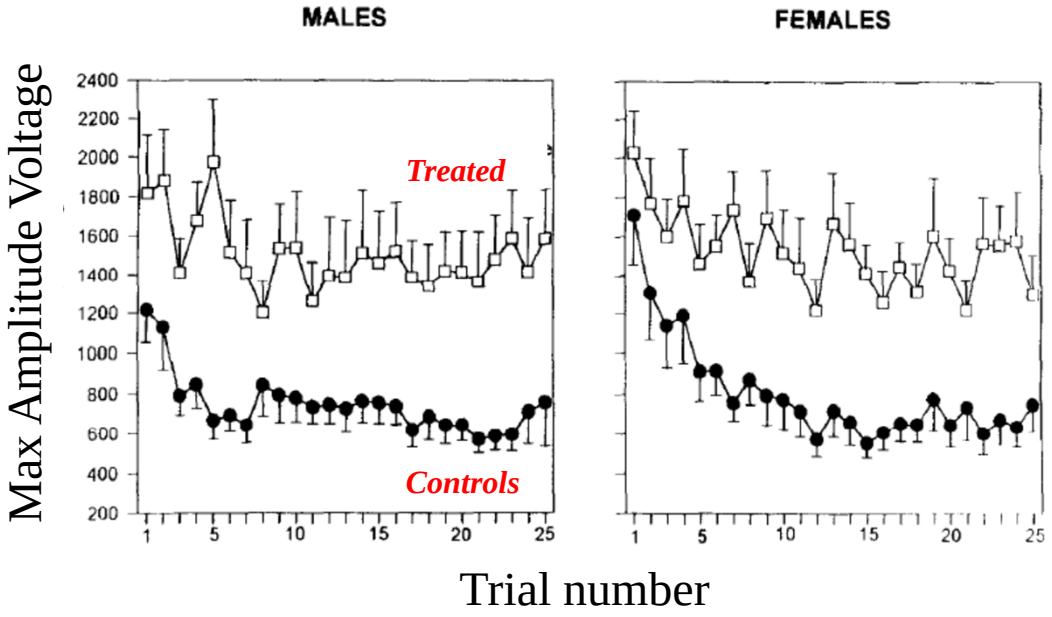
Effects from multiple chemicals show no clear correlation between treatment related decreases in body weight and decreases in ASR

If body weight is a covariate the data must be interpreted with caution

Cocaine and Maternal Weight



Methimazole and Adult Offspring Weight



Body weight: 15% decrease in weight gain of dams – no change in offspring

Startle: No changes in adult offspring

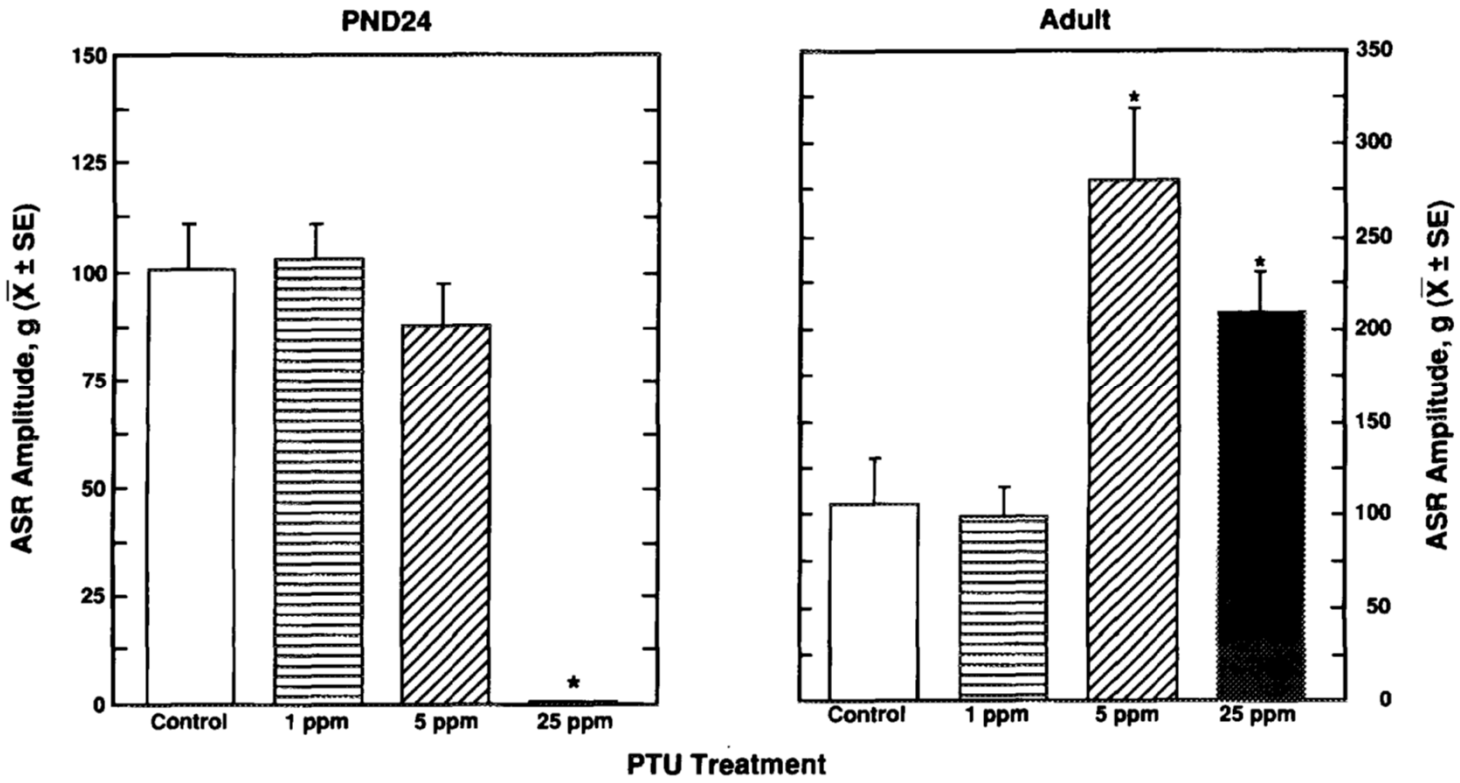
Foss and Riley, 1991

Body weight: Decreased in dam and offspring 12% males and 7% females

Startle: increases of ~75% in adult offspring

Henck et al., 1996

Propylthiouracil and Postnatal Weight



- **Body Weight**: no change in maternal weight. Offspring weight decreased up to 25% at high dose
- **Startle:** decreased amplitude at PND24 and increased at PND75

Goldey et al 1995

Reminder: Important Information in Study Protocol and/or Report

- Look for
  - Type of device used (and calibration)
  - Treatment balanced across time of day and test boxes
  - Good environmental control (e.g., animal handling, noise)
  - Is variability excessive
  - Training and experience of technical staff
  - Experimenter blind with respect to treatment
    - *(not as important with automated equipment compared to FOB)*
  - Historical and positive control data

Take Home Messages

- The test apparatus is important
  - How is calibrated done?
- It is critical to control of external variables to prevent excessive variation
- Sex effects – statistical effect or an opinion?
- Statistical vs biological significance


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
BOLZÏN
MERCI
BÏYAN
SHUKRIA
DANKSCHEEN
TASHAKKUR ATU
YAQHANYELAY
SUKSAMA
EKHMET
MEHRBANI
GRAZIE
MAAKE
KOMAPSUMNIDA
JUSPAXAR
GOZAIMASHITA
EFCHARISTO
TINGKI
SPASSIBO
SNACHALHUYA
NUHUN
CHALTU
WABEEJA
MAITEKA
YUSPAGARÄTAM
HUI
GUI
HATUR
UNALCHÉEESH
DHANYABAAD
ANIHA
ATTO
MERSI
SPASIBO
DENKAUJA
NENACHALHYA
EKOJU
SIKOMO
MAKETAU
MINMONCHAR
PALDIES
LAH
SANCO
MERASTAWHY
GAEJTHO
TAVTAPUCH
MEDAWAGSE
AGUYJE
FAKAAUE
BAIKA

References

Crofton KM, Makris SL, Sette WF, Mendez E, Raffaele KC. A qualitative retrospective analysis of positive control data in developmental neurotoxicity studies. Neurotoxicol Teratol. 2004 26(3):345-52.

Davis M. Sensitization of the rat startle response by noise. J Comp Physiol Psychol. 1974 Sep;87(3):571-81.

Davis M, Gendelman DS, Tischler MD, Gendelman PM. A primary acoustic startle circuit: lesion and stimulation studies. J Neurosci. 1982 Jun;2(6):791-805.

Dean, K. F., Sheets, L. P., Crofton, K. M., & Reiter, L. W. (1990). The effect of age and experience on inhibition of the acoustic startle response by gaps in background noise. *Psychobiology*, *18*(1), 89-95.

Davis, M. the Mammalian Startle Response, Chapter 10. In: Eaton, Robert C. (editor) Neural mechanisms of startle behavior. Springer Science & Business, ISBN 0-306-41556-9, 1984.

Eaton, Robert C. (editor) Neural mechanisms of startle behavior. Springer Science & Business Media, ISBN 0-306-41556-9, 1984

Foss JA, Riley EP. Failure of acute cocaine administration to differentially affect acoustic startle and activity in rats prenatally exposed to cocaine. Neurotox Teratol. 1991 3(5):547-51.

Glowa JR, Hansen CT. Differences in response to an acoustic startle stimulus among forty-six rat strains. Behav Genet. 1994 Jan;24(1):79-84.

Goldey ES, Kehn LS, Rehnberg GL, Crofton KM. Effects of developmental hypothyroidism on auditory and motor 08function in the rat. Toxicol Appl Pharmacol. 1995 135(1):67-76.

Henck JW, Frahm DT, Anderson JA. Validation of automated behavioral test systems. Neurotoxicol Teratol. 1996 Mar-Apr;18(2):189-97. doi: 10.1016/0892-0362(95)02058-6.

Hoffman HS, Ison JR. Reflex modification in the domain of startle: I. Some empirical findings and their implications for how the nervous system processes sensory input. Psychol Rev. 1980 Mar;87(2):175-89.

References (con't)

Hoffman HS, Ison JR. Reflex modification in the domain of startle: I. Some empirical findings and their implications for how the nervous system processes sensory input. Psychol Rev. 1980 Mar;87(2):175-89.

Koch M. The neurobiology of startle. Prog Neurobiol. 1999 Oct;59(2):107-28.

Marable BR, Maurissen JP. Validation of an auditory startle response system using chemicals or parametric modulation as positive controls. Neurotoxicol Teratol. 2004 26(2):231-7.

NAFTA (2016), North American Free Trade Agreement (NAFTA). Technical Working Group on Pesticides (TWG). Developmental Neurotoxicity Study Guidance Document. December, 2016. http://www.oecd-ilibrary.org/environment/test-no-426-developmental-neurotoxicity-study_9789264067394-en;jsessionid=7bt5uldknlfjn.x-oecd-live-01

OECD (2007), Test No. 426: Developmental Neurotoxicity Study (OECD Publishing).

OECD (2012), Organisation de coopération et de développement économiques. Test No. 443: Extended One-Generation Reproductive Toxicity Study. OECD Publishing.

Pantoni MM, Herrera GM, Van Alstyne KR, Anagnostaras SG. Quantifying the Acoustic Startle Response in Mice Using Standard Digital Video. Front Behav Neurosci. 2020 Jun 3;14:83.

Raffaele KC, Fisher JE Jr, Hancock S, Hazelden K, Sobrian SK. Determining normal variability in a developmental neurotoxicity test: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. Neurotox Teratol 2008 30(4):288-325.

Sheets LP, Dean KF, Reiter LW. Ontogeny of the acoustic startle response and sensitization to background noise in the rat. Behav Neurosci. 1988 Oct;102(5):706-13.

USEPA (1998), 'Health effects guidelines OPPTS 970.6300: Developmental Neurotoxicity Study. https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/series-870-health-effects-test-guidelines'.