**Analysis of CKC Tracking Test Records from 2005 to 2018**

**by Laura McKay and L. R. Schaeffer**

**March, 2022**

**Purpose of the Analysis**

To calculate pass rates for CKC tracking tests from 2005 to 2018 and to determine if and how much pass rates are affected by the year, the month (across years), the club and the judge.

**Raw Data**

Data were gathered from the archived Results of Tracking Tests published on the CKC website. All tracking tests reported between Jan. 1, 2005 and Dec. 31 2018 were included and comprised 628 tests. For each level (TD, TDX, UTD, UTDX) the data set included the number entered, the number passed, the date of the test, the judge and the club. Some tests were held over more than one day but were recorded by the CKC as one test.

**Data Structure**

**Table 1. Summary of the Number of Entries, Passes, Years, Months, Clubs and Judges by Level**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test Level | Number Entered | Number Passed | % Passed | Number of Years | Number of Months\* | Number of Clubs | Number of Judges |
| TD | 1991 | 1131 | 56.8 | 14 | 9 | 45 | 35 |
| TDX | 1159 | 407 | 35.1 | 14 | 9 | 45 | 35 |
| UTD | 782 | 223 | 28.5 | 14 | 10\* | 26 | 22 |
| UTDX | 457 | 92 | 20.1 | 13\*\* | 9 | 21 | 19 |
| Total | 4389 | 1853 |  |  |  |  |  |

\* The first UTD test was held Jan. 1, 2005, but all other tests were held in the months from March to November.

\*\* The first UTDX test was held in 2006.

**Entries by Year:**

Total entries for all levels rose from a low of 258 in 2005 to peak at 387 in 2011 (Figure 1). The rise in entries may be due to the addition and proliferation of UTD and then UTDX tests in 2005 and 2006. After 2011 total entries dropped and then remained relatively constant from 2012 to 2018 between 274 and 311 entries each year.

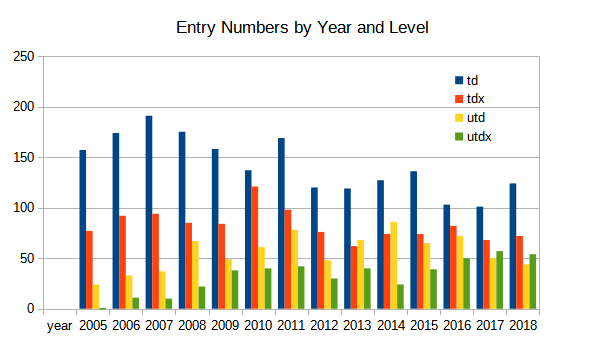
A preliminary scan of the entries each year shows fluctuations from year to year for all levels (Figure 2). There are more TD entries in each year than entries at any other level. TD entries peaked in 2007, with an apparent decline afterwards. This may indicate that more TD candidates were passingon their first attempt in later years.

TDX entries were relatively stable over the time period, peaking in 2011.

The entries for urban tests increased over the first few years but appeared to level off after that.

**Figure 1.**

**Figure 2.**



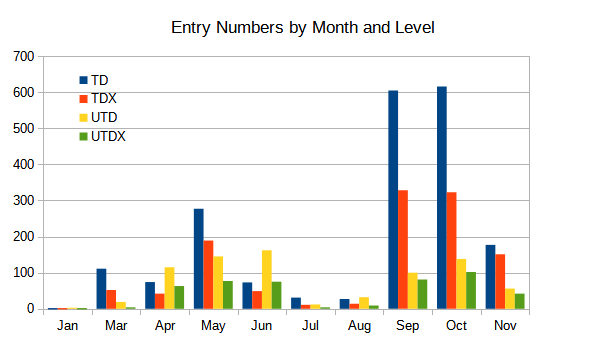
**Entry Numbers by Month:**

Total entries were highest in October and September with a smaller Spring peak in May (Figure 3). July and August had few entries some of which were associated with National Specialty Shows.

**Figure 3.**

There appear to be more entries in field tests in the Fall, particularly in September and October compared with the Spring (see Figure 4). Urban entries are more even across the seasons.

**Figure 4.**



**Entry Numbers by Club**

The distribution of entries across clubs is quite uneven with a handful of clubs having the bulk of the entries over the time span of the data (Table 2). The clubs with the highest entries are primarily Ontario clubs except for CSDCA in Alberta and FPOTC in BC. Within the Ontario clubs, three are located in southern Ontario (CCTC, NTTA and GSDCL). All the clubs with high entry numbers have been holding tests most years throughout the period from 2005 to 2018.

For field tracking tests, the high-entry clubs had 46% (TD) and 48% (TDX) of all the entries. For urban tracking tests, there were just four clubs that together had 50% of all entries. Three are Ontario clubs (CCTC, EPSTA and TBKTC) and one is in BC (FPOTC). The three Ontario clubs serve different areas: the south, the near north and the north.

At each level, there were several or even many clubs that had mid-range entry numbers and several that had very few entries (less than 10 each) over the time period. Half of all clubs holding tests had total entries of less than 23, 15, 17 and 18 dogs for TD, TDX, UTD and UTDX, respectively (Tables 1 and 2). Clubs with less than 10 entries over the time period represented 18%, 40%, 27% and 43% of all clubs holding TD, TDX, UTD and UTDX tests, respectively. However, these low-entry clubs accounted for a very small percent of the total entry for TD and UTD and a somewhat higher proportion of the entry for TDX and UTDX (Table 2). Thus the distributions of clubs with respect to the number of entries was more skewed towards low numbers for the higher levels of urban and field tests compared with the entry levels. Some of the clubs with lower entries were active earlier and stopped holding tests or started holding tests part-way through the period studied. Other clubs typically hold only one test a year and or tend to have small entry numbers per test. Clubs holding tests at National Specialties may not have held tests every year and or entries may have been low in each test that was held.

**Table 2. Entry Numbers for Clubs with Highest and Lowest Entries**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | TD | | | TDX | | UTD | | UTDX | |
|  | Club | | Number of Entries | Club | Number of Entries | Club | Number of Entries | Club | Number of Entries |
| Highest Entry | CCTC | | 219 | CSDCA | 126 | FPOTC | 114 | EPSTA | 67 |
| CSDCA | | 186 | EPSTA | 110 | EPSTA | 106 | TBKTC | 60 |
| FPOTC | | 115 | TBKTC | 101 | TBKTC | 89 | CCTC | 55 |
| NTTA | | 105 | CCTC | 80 | CCTC | 82 | FPOTC | 46 |
| EPSTA | | 98 | FPOTC | 65 |  |  |  |  |
| TBKTC | | 98 | NTTA | 59 |  |  |  |  |
| GSDCL | | 90 |  |  |  |  |  |  |
| Total of Highest Entry Clubs | | 911 | |  | 541 |  | 391 |  | 228 |
| % of total entry | | 45.8 | |  | 47.8 |  | 50.0 |  | 49.9 |
|  | | | | | | | | | |
| Lowest Entry (less than 10 dogs) | 8 clubs pooled | | 36 | 18 clubs pooled | 127 | 7 clubs pooled | 21 | 9 clubs pooled | 48 |
| % of total entry | | 1.8 | |  | 11.0 |  | 2.9 |  | 9.2 |

CCTC=Cross Country Tracking Club (ON)

CSDCA=Competitive Service Dog Club of Alberta (AB)

EPSTA = EPS Training Associates (ON)

FPOTC=Forbidden Plateau Obedience & Tracking Club (BC)

GSDCL=German Shepherd Dog Club of London (ON)

NTTA= Newark Training & Tracking Associates (ON)

TBKTC=Thunder Bay Kennel & Training Club (ON)

**Entry Numbers by Judge:**

As with clubs, there were a handful of judges that judged a high percentage of the dogs entered at each level (Table 3). These judges tested between 50 and 62% of all dogs entered depending on the level. All of the urban judges with the highest entry were field judges that were grandfathered in as urban judges in 2005.

In contrast, judges that judged less than 10 dogs over this time period judged a low percentage of all dogs at each level. These minimally active judges represent 11%, 28%, 27% and 37% of all judges, for TD, TDX, UTD and UTDX, respectively (Tables 1 and 3). Half of all judges active during this period judged less than 20, 19, 22 and 20 dogs in TD, TDX, UTD and UTDX, respectively. Many of the less active judges were American judges, new judges or retiring judges. For TD, the remaining judges (more than 10 dogs but less than the high-entry judges) were relatively equally distributed over the range of the number of entries. For TDX, the remaining judges judged less than 40 dogs each. The distribution of judges with respect to number of dogs judged was even more skewed than the distribution of clubs. For UTD and UTDX, the remaining judges were similarly skewed.

M-P Babin has judged more dogs combined over all the levels than any other judge (636).

**Table 3. Entry Numbers for Judges with Highest and Lowest Entries**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | TD | | TDX | | UTD | | UTDX | |
|  | Judge | No. of Entries | Judge | No. of Entries | Judge | No. of Entries | Judge | No. of Entries |
| Highest Entry | G Martin | 261 | G Martin | 182 | J Mcleod | 176 | M-P Babin | 99 |
| M-P Babin | 254 | M-P Babin | 167 | M-P Babin | 116 | J Mcleod | 66 |
| J Book | 129 | J Book | 72 | J Wilhelm | 68 | J Wilhelm | 39 |
| J Mcleod | 120 | J Mcleod | 63+ | S Smith | 65 | J Book | 35 |
| J Wilhelm | 105 | L Weaver | 53 | J Book | 62 | S Smith | 32 |
| L Weaver | 98 | J Wilhelm | 50 |  |  |  |  |
| S Smith | 89 |  |  |  |  |  |  |
| S Sorenson | 89 |  |  |  |  |  |  |
| Total of highest entry judges | | 1145 |  | 587 |  | 487 |  | 271 |
| % of total entry | | 57.5 |  | 50.6 |  | 62.3 |  | 59.3 |
|  | | | | | | | | |
| Lowest Entry  (less than 10 dogs) | 4 judges\* | 26 | 10 judges\* | 52 | 6 judges | 23 | 7 Judges | 42 |
| % of total entry | | 1.3 |  | 4.5 |  | 2.9 |  | 9.2 |

+ J Mcleod also judged 11 TDX dogs split with G Roe which are not included because the exact number that each judged is not known.

\* excludes one unknown judge that judged a test on Mar 6, 2005 since the judge likely judged other dogs as well.

**Analysis**

The raw data were the number entered and number passed so records were generated for each dog entered as 1 or 0 for pass/fail plus level, year, month, club and judge. The generated data were analyzed with R software (see the R home page http://.R-project.org) using a separate least squares analysis (AOV) for each level. The model included year, month, club and judge effects and a residual error. Each observation is modelled as a combination of a year effect, a month effect, a club effect and a judge effect. Thus the solution for each year accounts for differences in the months, the clubs or the judges and likewise for the month solutions, the club solutions and the judge solutions.

**Results of the Analysis:**

The results of the analysis are presented in Table 4. R-squared is a measure of how well the model explains the data, ranging from 0 (doesn’t explain the data at all) to 1 (explains the data exactly). Although the model explains some of the differences in pass rates, there is still a lot that is not explained by the model. The model doesn’t explicitly include effects due to the weather (temperature, wind, rain, etc.), the dog/handler (age of dog, breed of dog, training level, experience of the handler, etc.) or various other factors that may influence pass rates. Information about these other factors was not available in the data. However, some effects of weather may be included in the Month and Club effects, since weather patterns may tend to differ from one month to another and from one geographical area to another. Dog/handler abilities may vary due to geographical area and therefore, Club effects may reflect some of the differences in pass rate due to the dog and or handler. Month and Club may be partly confounded because many clubs hold their tests at the same time each year. Regardless, there would still be random effects due to weather, dog/handler and other factors.

The significance level measures the likelihood that the effect of the factor is real. The P value measures the statistical probability that the effect of the factor is due to chance alone, rather than due to a real effect. Hence, a very low P value for an effect implies that the effect of that factor is real (i.e. significant). Typically, a P value of less than 0.1 (\*) is considered marginally significant; a P value less than 0.01 (\*\*) is considered significant; and a P value less than 0.001 (\*\*\*) is considered highly significant.

**Table 4. Significance of the Factors in the Model by Level**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test Level | R-squared | Significance Level (P value) for the Effect of | | | | |
| Model | Year | Month | Club | Judge |
| TD | 0.1174 | 0.000\*\*\* | 0.281 NS | 0.002\*\*\* | 0.000\*\*\* | 0.001\*\*\* |
| TDX | 0.1418 | 0.000\*\*\* | 0.434 NS | 0.168 NS | 0.020\* | 0.134 NS |
| UTD | 0.1550 | 0.000\*\*\* | 0.072\* | 0.000\*\*\* | 0.001\*\*\* | 0.003\*\* |
| UTDX | 0.2776 | 0.000\*\*\* | 0.072\* | 0.919 NS | 0.000\*\*\* | 0.000\*\*\* |

Solutions for each factor were generated in the analysis. The solutions reflect the effect of that factor has on pass rates, corrected for all the other factors in the model. For example, a solution for 2010 is the effect that year had on the pass rate independent of the month, club and judge effects; a solution for May is the effect of that month independent of the year, the club or the judge; a solution for CCTC is independent of the year, month and judge effects; and a solution for M-P Babin is independent of when or where she judged.

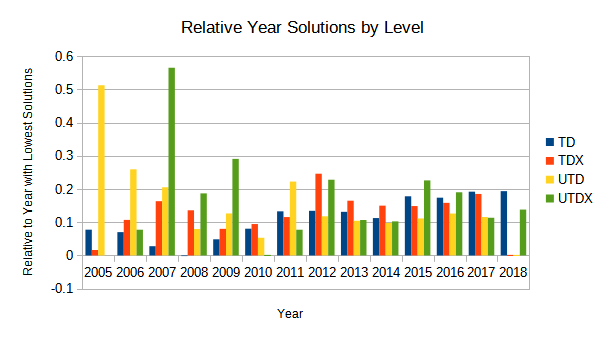
**Relative Year Solutions**

The solutions for each year for every level were adjusted so that the year with the lowest solution within each level was set to zero and all other levels are plotted relative to the lowest. We did this for ease of comparison visually between solutions for years within levels. Note that the actual solutions ranged from positive to negative, but absolute differences between pairs of solutions were preserved.

The relative solutions for field tests (TD and TDX) do not appear to vary greatly over the years (Figure 5) which is reflected in the lack of significance of the Year effect in the Model for TD and TDX (Table 4).

For urban test levels (UTD and UTDX) the relative Year solutions for 2005 for UTD and 2007 for UTDX stand out as much higher than the adjusted solutions for other years, which may explain the marginal significance of Year effect for these two levels (Figure 5 and Table 4). However, both of these data points result from a small number of entries and are likely outliers. The urban tests were introduced in 2005 (UTD) and 2006 (UTDX) and the relative year solutions appear to fluctuate initially, but become more stable by 2012. Any apparent instability in the Year effects during the first few years after the introduction of the urban tests may be due to judges, handlers and dogs adjusting to the unique demands of urban tests.

**Figure 5.**



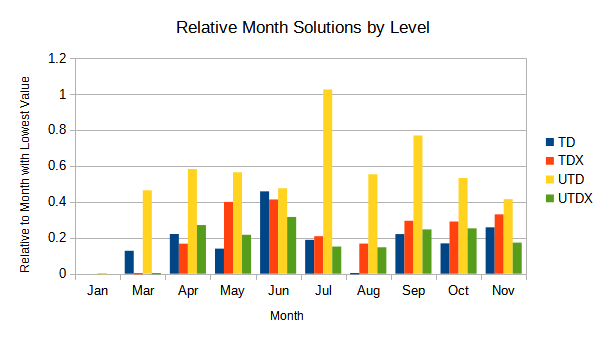
**Relative Month Solutions**

As for Year solutions, Month solutions were adjusted relative to the month with the lowest value within each test level (Figure 6). Month was significant for TD and UTD (Table 4). For the TD level, June was the best month and August was the poorest. Both months have low entry numbers (Figure 4), so both solutions may be the result of random variation. Factors that may alter pass rates for TD include vegetation heights and lushness in fields, temperature and sufficient time to train prior to the test.

For UTD, July was the best month and January was the poorest (Figure 6). Both these months had very low entry numbers, so these two months were likely outliers. Apart from these months with low entry numbers, solutions for pass rate peak in April/May and Sept/Oct. These peaks may be related to temperatures and or opportunities to train prior to the test.

For TDX and UTDX, the Month effect was not significant (Table 4) and the solutions did not vary month to month as much as they did for TD and UTD (Figure 6).

**Figure 6.**



**Club effects:**

Club effects were highly significant for all levels except TDX (Table 4). Differences in pass rate among clubs may result from significant differences in sites, weather and dog/handler abilities. Some clubs may have sites that have less contamination and or less extreme temperatures, wind or rain. Their entrants may be more familiar with the sites and weather than clubs that the draw entrants from a wider area. Some clubs may hold tests in months that are more favourable.

Club effects were just marginally significant for TDX (Table 4). Possibly site and weather differences are smaller among clubs or are less important for TDX dogs and handlers.

**Judge Effects:**

Judge effects were highly significant for all levels except TDX (Table 4). Judge effects should be relatively free of site, weather and dog/handler abilities. Judge effects may be due to differences in how judges plot tracks and or how they judge while the dog is tracking. Judges may differ in their perception of what constitutes a fair track or a valid track for a given level. Judges that judge primarily on sites that they are familiar with may have higher pass rates than judges who have primarily judged on sites they are unfamiliar with, similar to a home-vs-away effect on sports teams’ success rates in games.

**Summary:**

1. The raw pass rates over the time period from 2005 to 2018 were 57%, 35%, 29% and 20% for TD, TDX, UTD and UTDX, respectively.

2. Entry numbers for field tests (TD and TDX) have been relatively stable over the years in this time period. Urban tests (UTD and UTDX) took about 4 years from the year of the first test (2005 for UTD and 2006 for UTDX) to build up to a level after which entry numbers were relatively stable.

3. Most field tracking tests (TD and TDX) were held in September and October, with a smaller peak in May. Urban tracking tests (UTD and UTDX) were spread out more over the year.

4. At each level, a handful of clubs had 46% to 50% of all entries. These clubs ran tests in all or most years from 2005 to 2018. Of the remaining clubs, the distribution over the number of entries was skewed towards low numbers, particularly for TDX, UTD and UTDX so that large numbers of clubs had low numbers of entries.

5. At each level, a handful of judges tested 50% to 62% of all dogs over the time period. Most of these judges were active throughout the time period. The five urban judges that judged the most dogs were all grandfathered in as urban judges in 2005. Of the remaining judges, the distribution over the number of dogs judged was skewed towards low numbers, particularly for TDX, UTD and UTDX.

6. The data were analyzed with a Least Squares model that included Year, Month, Club and Judge effects. Year did not significantly affect pass rates for TD and TDX and was at most marginally significant for pass rates for UTD and UTDX. Month influenced pass rates for TD and UTD, but not for TDX or UTDX. Club influenced pass rates significantly for all levels except TDX, where it was marginally significant. Judge influenced pass rates significantly for all levels except for TDX. Although there were significant effects for several of the factors in the model, the model did not explain a high proportion of the differences in pass rate at any of the levels. Random factors such as weather, site variations and dog/handler effects that cannot be identified in the data set produce the majority of differences in pass rates.